

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2001, 15:47:42 ; Search time 77.49 Seconds
(without alignments)
6.259 Million cell updates/sec

Title: US-09-630-345-1

Perfect score: 44

Sequence: 1 LMPTYLK 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

al number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- A_Geneseq_0601.*
- 1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.*
 - 2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.*
 - 3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.*
 - 4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.*
 - 5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.*
 - 6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.*
 - 7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.*
 - 8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.*
 - 9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.*
 - 10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.*
 - 11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.*
 - 12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.*
 - 13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.*
 - 14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.*
 - 15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.*
 - 16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.*
 - 17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.*
 - 18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.*
 - 19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.*
 - 20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.*
 - 21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.*
 - 22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	44	100.0	8	19 AAW50937	Vasoactive intesti
2	44	100.0	8	21 AAB08298	Vasoactive intesti
3	44	100.0	8	21 AAB08306	Amino acid sequenc
4	39	88.6	8	21 AAB08305	Amino acid sequenc
5	35	79.5	8	21 AAB08304	Amino acid sequenc
6	35	79.5	450	21 AAB63104	Human secreted pro
7	35	79.5	483	20 AAW93163	Human MMP-20 prote
8	33	75.0	261	19 AAW76253	Human matrilysin-1
9	33	75.0	462	22 AAB49983	Murine macrophage
10	32	72.7	173	22 AAB74618	Human 72 kDa gelat
11	32	72.7	173	22 AAB66862	Matrilysin catalyt

12	32	72.7	264	19 AAW76254	Human matrilysin p
13	32	72.7	271	17 AAR96211	Human recombinant
14	32	72.7	297	21 AAB43585	Human cancer assoc
15	32	72.7	370	16 AAB87016	Type I matrix meta
16	32	72.7	409	21 AAB26439	Drosophila melanog
17	32	72.7	807	21 AAY57288	Human GPCR protein
18	32	72.7	986	21 AAY40440	Human brain-derive
19	31	70.5	8	20 AAW75457	Mammalian tub prot
20	31	70.5	12	20 AAW97073	Peptidomimetic cap
21	31	70.5	58	20 AAY12222	Human 5' EST seque
22	31	70.5	114	21 AAG17440	Arabidopsis thalia
23	31	70.5	146	21 AAG17439	Arabidopsis thalia
24	31	70.5	163	21 AAG44201	Arabidopsis thalia
25	31	70.5	255	21 AAG44200	Arabidopsis thalia
26	31	70.5	287	21 AAG44199	Arabidopsis thalia
27	31	70.5	325	14 AAR42750	ced-4 gene II prod
28	31	70.5	457	10 AAP93628	Sequence of human
29	31	70.5	459	18 AAB36485	Mouse TUB Form I.
30	31	70.5	459	21 AAB26400	Mouse tub Form I
31	31	70.5	460	18 AAB36488	Human TUB Form 6.
32	31	70.5	460	21 AAB26903	Human TUB Form 6 p
33	31	70.5	469	8 AAP70611	Sequence encoded b
34	31	70.5	469	21 AAB10655	BPV1 L1 fusion pro
35	31	70.5	469	21 AAB10660	BPV2 L1 fusion pro
36	31	70.5	495	6 AAP51099	Sequence of bovine
37	31	70.5	495	20 AAW88482	Bovine papillomavi
38	31	70.5	505	18 AAW36486	Mouse TUB Form II.
39	31	70.5	505	18 AAW10728	Mouse tub gene pro
40	31	70.5	505	19 AAW54367	Mouse tub polypept
41	31	70.5	505	20 AAW75450	Mouse wild type tu
42	31	70.5	505	21 AAB26901	Mouse tub Form II
43	31	70.5	506	18 AAW36497	Human TUB Form 4.
44	31	70.5	506	18 AAW10729	Human tub gene pro
45	31	70.5	506	19 AAW54368	Human tub polypept

ALIGNMENTS

RESULT 1

AAW50937
ID AAW50937 standard; peptide; 8 AA.

XX AC AAW50937;

XX DT 31-JUL-1998 (first entry)

XX DE Vasoactive intestinal peptide receptor binding inhibitor (VIP2).

XX Vasoactive intestinal peptide; VIP; antagonist; somatostatin; bombesin;

XX KW Substance P; cancer; inhibition.

XX OS Synthetic.

XX PN EP835662-A2.

XX PD 15-APR-1998.

XX PF 11-DEC-1996; 96EP-0309012.

XX PR 08-OCT-1996; 96US-0727679.

XX PR 16-AUG-1996; 96IN-0001822.

XX (NATIM-) NAT INST IMMUNOLOGY.

XX Jaggi M, Mukherjee R;

XX WPI; 1998-208959/19.

XX Composition containing analogues of vasoactive intestinal peptide,
PT somatostatin - bombesin and substance P, for treatment of tumours
PT and for inhibiting over-expression of these peptide(s)

XX

PS Claim 1; Page 4; 49pp; English.

XX The invention relates to a new composition which comprises: (i) the
CC somatostatin analogue SOM2 AGCKNPRDWTPTNSDC (3-14 disulphide bridge),
CC and (ii) at least 4 of the peptides: antagonist of vasoactive
CC intestinal peptide (VIP1); VIP receptor-binding inhibitor (VIP2); VIP
CC receptor antagonist (VIP3); somatostatin analogue (SOM1); bombesin
CC antagonist (BOM1) and substance P antagonist (SP1). Also claimed are
CC more general compositions containing peptide analogues of somatostatin,
CC VIP, bombesin and substance P. The compositions are used in human or
CC veterinary medicine: (a) to kill (or inhibit multiplication of) tumour
CC or cancer cells, particularly for treatment of leukaemia, lymphoma,
CC adenocarcinoma of stomach, pancreas or prostate, or cancer of lung,
CC breast, kidney or particularly rectum and colon, and (b) to prevent,
CC inhibit or modulate over-expression of, e.g. VIP. A wide range of cancer
CC cells express receptors for VIP, somatostatin, bombesin and/or substance
CC P. The present sequence represents VIP receptor-binding inhibitor
CC (VIP2).

XX Sequence 8 AA:

Query Match 100.0%; Score 44; DB 19; Length 8;
Best Local Similarity 100.0%; Pred. No. 3.4e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMYPITYLK 8
Db 1 LMYPITYLK 8

RESULT 2

AA08298 AAB08298 standard; peptide; 8 AA.

XX Vasoactive intestinal peptide (VIP) analogue VIP2.
XX Vasoactive intestinal peptide; VIP; analogue; somatostatin; SOM1; SOM2;
XX VIP1; VIP2; VIP3; BOM1; bombesin; SP1; substance P; MuJ-7; tumour growth;
XX tumour angiogenesis; metastasis; cancer; angiogenesis; adenocarcinoma;
XX leukaemia; lymphoma.
XX Synthetic.

WO200047221-A1.

17-AUG-2000.

11-FEB-2000; 2000WO-US03559.

11-FEB-1999; 99US-0248381.

(NAIM-) NAT INST IMMUNOLOGY.
(DABU-) DABUR RES FOUND.
(CORD/) CORD J I.

Mukherjee R, Jaggi M, Prasad S, Burman AC, Rajendran P, Mathur A;
Singh AT;

WPI; 2000-549083/50.

Novel therapeutically active composition comprising at least 5
PT peptides, useful for treating angiogenesis especially as a result of
XX adenocarcinomas -

PS Disclosure; Page 8; 42pp; English.

XX The present sequence represents an analogue of vasoactive intestinal
CC peptide (VIP). The specification describes therapeutically active

CC compositions comprising at least one analogue of somatostatin (chosen
CC from SOM1 and SOM2) and at least four analogues chosen from VIP1 (a
CC VIP antagonist), VIP2 (a VIP receptor binding inhibitor), VIP3 (a VIP
CC receptor antagonist), BOM1 (a bombesin antagonist), and SP1 (a substance
CC P antagonist). The combination of these 7 analogues is known as MuJ-7.
CC MuJ-7 is used as an anticancer drug to restrict tumour growth and spread
CC by inhibiting tumour angiogenesis. MuJ-7, in addition, inhibits
CC metastasis through its antiangiogenic activity in all cancers. The
CC peptides are useful for the treatment and prevention of angiogenesis,
CC especially as a result of adenocarcinomas of the colon, breast, lung,
CC prostate, kidney, leukemias or lymphomas.

XX Sequence 8 AA:

Query Match 100.0%; Score 44; DB 21; Length 8;
Best Local Similarity 100.0%; Pred. No. 3.4e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMYPITYLK 8
Db 1 LMYPITYLK 8

RESULT 3

AA08306 AAB08306 standard; peptide; 8 AA.

XX Vasoactive intestinal peptide; VIP; analogue; somatostatin; SOM1; SOM2;
XX VIP1; VIP2; VIP3; BOM1; bombesin; SP1; substance P; MuJ-7; tumour growth;
XX tumour angiogenesis; metastasis; cancer; angiogenesis; adenocarcinoma;
XX leukaemia; lymphoma.
XX Synthetic.

WO200047221-A1.

17-AUG-2000.

11-FEB-2000; 2000WO-US03559.

11-FEB-1999; 99US-0248381.

(NAIM-) NAT INST IMMUNOLOGY.
(DABU-) DABUR RES FOUND.
(CORD/) CORD J I.

Mukherjee R, Jaggi M, Prasad S, Burman AC, Rajendran P, Mathur A;
Singh AT;

WPI; 2000-549083/50.

Novel therapeutically active composition comprising at least 5
PT peptides, useful for treating angiogenesis especially as a result of
XX adenocarcinomas -

PS Claim 11; Page 31; 42pp; English.

XX AAB08304-15 represent peptides which have an antiangiogenic effect. The
CC specification describes therapeutically active compositions comprising
CC at least one analogue of somatostatin (chosen from SOM1 and SOM2), and
CC at least four analogues chosen from vasoactive intestinal peptide (VIP)
CC 1 (a VIP antagonist), VIP2 (a VIP receptor binding inhibitor), VIP3 (a
CC VIP receptor antagonist), BOM1 (a bombesin antagonist), and SPI (a
CC substance P antagonist). The combination of these 7 analogues is known as
CC MuJ-7. MuJ-7 is used as an anticancer drug to restrict tumour growth and
CC spread by inhibiting tumour angiogenesis. MuJ-7, in addition, inhibits
CC metastasis through its antiangiogenic activity in all cancers. The
CC peptides are useful for the treatment and prevention of angiogenesis,
CC especially as a result of adenocarcinomas of the colon, breast, lung,
CC prostate, kidney, leukemias or lymphomas.
XX SQ Sequence 8 AA;

Query Match 100.0%; Score 44; DB 21; Length 8;
Best Local Similarity 100.0%; Pred. No. 3.4e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LMYPYLYLK 8
| | | | | | |
DB 1 lmyptylk 8

RESULT 4
AAB08305
ID AAB08305 standard; peptide; 8 AA.
XX AC AAB08305;
XX DT 04-DEC-2000 (first entry)
XX DE Amino acid sequence of an antiangiogenic peptide.
XX KW Vasoactive intestinal peptide; VIP; analogue; somatostatin; SOM1; SOM2;
KW VIP1; VIP2; VIP3; BOM1; bombesin; SPI; substance P; MuJ-7; tumour growth;
KW tumour angiogenesis; metastasis; cancer; angiogenesis; adenocarcinoma;
KW leukaemia; lymphoma.
XX OS Synthetic.
XX FH Key Location/Qualifiers
FT Misc-difference 1 /note= "D-form residue"
FT Modified-site 7 /label= Aib, Deg
FT /note= "alpha-aminoisobutyric acid or alpha, alpha
diethyl glycine"
XX WO200047221-A1.
XX PN 17-AUG-2000.
XX PF 11-FEB-2000; 2000WO-US03559.
XX PR 11-FEB-1999; 99US-0248381.
XX (NAIM-) NAT INST IMMUNOLOGY.
XX (DABU-) DABUR RES FOUND.
XX PA (CORD/) CORD J I.
XX PI Mukherjee R, Jaggi M, Prasad S, Burman AC, Rajendran P, Mathur A;
XX PI Singh AT;
XX DR WPI; 2000-549083/50.
XX Novel therapeutically active composition comprising at least 5
PT peptides, useful for treating angiogenesis especially as a result of
PT adenocarcinomas -
XX

PS Claim 11; Page 31; 42pp; English.
XX AAB08304-15 represent peptides which have an antiangiogenic effect. The
CC specification describes therapeutically active compositions comprising
CC at least one analogue of somatostatin (chosen from SOM1 and SOM2), and
CC at least four analogues chosen from vasoactive intestinal peptide (VIP)
CC 1 (a VIP antagonist), VIP2 (a VIP receptor binding inhibitor), VIP3 (a
CC VIP receptor antagonist), BOM1 (a bombesin antagonist), and SPI (a
CC substance P antagonist). The combination of these 7 analogues is known as
CC MuJ-7. MuJ-7 is used as an anticancer drug to restrict tumour growth and
CC spread by inhibiting tumour angiogenesis. MuJ-7, in addition, inhibits
CC metastasis through its antiangiogenic activity in all cancers. The
CC peptides are useful for the treatment and prevention of angiogenesis,
CC especially as a result of adenocarcinomas of the colon, breast, lung,
CC prostate, kidney, leukemias or lymphomas.
XX SQ Sequence 8 AA;

Query Match 88.6%; Score 39; DB 21; Length 8;
Best Local Similarity 87.5%; Pred. No. 3.4e+05;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 LMYPYLYLK 8
| | | | | | |
DB 1 lmyptyxx 8

RESULT 5
AAB08304
ID AAB08304 standard; peptide; 8 AA.
XX AC AAB08304;
XX DT 04-DEC-2000 (first entry)
XX DE Amino acid sequence of an antiangiogenic peptide.
XX KW Vasoactive intestinal peptide; VIP; analogue; somatostatin; SOM1; SOM2;
KW VIP1; VIP2; VIP3; BOM1; bombesin; SPI; substance P; MuJ-7; tumour growth;
KW tumour angiogenesis; metastasis; cancer; angiogenesis; adenocarcinoma;
KW leukaemia; lymphoma.
XX OS Synthetic.
XX FH Key Location/Qualifiers
FT Modified-site 1 /label= Aib
FT /note= "alpha-aminoisobutyric acid"
FT Modified-site 7 /label= Aib, Deg
FT /note= "alpha-aminoisobutyric acid or alpha, alpha
diethyl glycine"
XX WO200047221-A1.
XX PN 17-AUG-2000.
XX PF 11-FEB-2000; 2000WO-US03559.
XX PR 11-FEB-1999; 99US-0248381.
XX (NAIM-) NAT INST IMMUNOLOGY.
XX (DABU-) DABUR RES FOUND.
XX PA (CORD/) CORD J I.
XX PI Mukherjee R, Jaggi M, Prasad S, Burman AC, Rajendran P, Mathur A;
XX PI Singh AT;
XX DR WPI; 2000-549083/50.
XX Novel therapeutically active composition comprising at least 5
PT peptides, useful for treating angiogenesis especially as a result of
PT

PT	adenocarcinomas -
XX	
XX	Claim 18; Page 36; 42pp; English.
XX	
XX	AAB08304-15 represent peptides which have an antiangiogenic effect. The specification describes therapeutically active compositions comprising at least one analogue of somatostatin (chosen from SOM1 and SOM2), and at least four analogues chosen from vasoactive intestinal peptide (VIP) 1 (a VIP antagonist), VIP2 (a VIP receptor binding inhibitor), VIP3 (a VIP receptor antagonist), BOM1 (a bombesin antagonist), and SPl (a substance P antagonist). The combination of these 7 analogues is known as MuJ-7. MuJ-7 is used as an anticancer drug to restrict tumour growth and spread by inhibiting tumour angiogenesis. MuJ-7, in addition, inhibits metastasis through its antiangiogenic activity in all cancers. The peptides are useful for the treatment and prevention of angiogenesis, especially as a result of adenocarcinomas of the colon, breast, lung, prostate, kidney, leukemias or lymphomas.
XX	Sequence 8 AA:
XX	

```
Query Match      79.5%; Score 35; DB 21; Length 8;
Best Local Similarity 85.7%; Pred. No. 3.4e+05;
Matches . 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

Qy	2 MYPTLK 8
Db	2 MYPTLYK 8

RESULT	6
AA0663104	
ID	AA0663104 standard; Protein: 450 AA.
XX	
AC	AA0663104:
XX	
DT	26-MAR-2001 (first entry)
DE	Human secreted protein sequence encoded by gene 18 SEQ ID NO:114.
XX	
XX	Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;
XX	antirheumatic; antiproliferative; cytostatic; cardiac; vasotropic;
KW	cerebroprotective; neurotropic; neuroprotective; antibacterial; virucide;
KW	fungicide; ophthalmological; vulnerary; gene therapy; neoplasm;
KW	autoimmune disease; rheumatoid arthritis; hyperproliferative disorder;
KW	cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW	cerebral ischaemia; angiogenesis; nervous system disorder; infection;
KW	Alzheimer's disease; ocular disorder; corneal infection; wound healing;
XX	skin aging; food additive; preservative.

XX	Homo sapiens.
OS	
XX	
XX	WO2000061748-A1.
PN	
XX	
XX	19-OCT-2000.
PD	
PD	
XX	
XX	06-APR-2000; 2000WO-US08982.
PF	
XX	
XX	
PR	09-APR-1999; 90US-0128696.
PR	14-JAN-2000; 2000US-0176069.
PR	
XX	
XX	(HUMA-) HUMAN GENOME SCI INC.
PA	
XX	
XX	
PI	Rosen CA, Ruben SM, Komatsoulis G;
XX	
XX	
DR	WPI: 2000-638566/61.

AAF22316 to AAF22363 encode the human secreted proteins given in AAB63049 to AAB63096. AAB63097 to AAB63132 represent more human secreted proteins and polypeptides homologous to them. Human secreted proteins have activities based on the tissues and cells the genes are expressed in. Examples of activities include: immunosuppressive; antiarthritic; antirheumatic; antiproliferative; cytostatic; cardiac; vasotropic; cerebroprotective; neurotropic; neuroprotective; antibacterial; virucide; fungicide; opthalmological; and vulnerary. The polynucleotides and proteins can be used to prevent, treat or ameliorate a medical condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used in diagnosing a pathological condition or susceptibility to a pathological condition. Disorders which are diagnosed or treated include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g. Alzheimer's disease, infections caused by bacteria, viruses and fungi and ocular disorders e.g. corneal infection. The polypeptides can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues and in chemotaxis. The polypeptides can also be used as a food additive or preservative to increase or decrease storage capabilities, fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors and other nutritional components. AAF22307 to AAF22315 and AAB63048 represent sequences used in the exemplification of the present invention.

Query Match 79.5%; Score 35; DB 21; Length 450;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 MYPTYL 7
pb 106 myptyl 111

RESULT	7
AAW93163	
ID	AAW93163 standard; Protein; 483 AA.
XX	
AC	AAW93163;
XX	
DT	21-MAY-1999 (first entry)
XX	
DE	Human MMP-20 protein.
XX	
KW	MMP-20; human; enamelysin; matrix metalloproteinase; teeth development;
KX	
KW	enamel; decomposition; maturation; matrix; mineralisation; drug.
XX	
OS	Homo sapiens.
XX	
PN	WO9906555-Al.
XX	
PD	11-FEB-1999.
XX	
PF	29-JUL-1998; 98WO-JP03381.
XX	
PR	29-JUL-1997; 97JP-0217038.
XX	
PA	(FUJY) FUJI YAKUHIIN KOGYO KK.
XX	
PI	Aoki T, Cuadra EL, Iwata K, Lopez-Otin C, Pendas AM;
XX	
P1	WPI: 1999-153785/13.
DR	N-PSDB; AAX22532.
XX	
PT	Dental pulp-originated matrix metalloproteinase (MMP)-20 - used for
PT	studies of development of teeth, enamel maturation process, and
PT	medical, and clinical applications

PS Claim 3; Page 89-90; 97pp; Japanese.

XX This sequence represents a novel human enamelysin (MMP-20) protein, a
 CC matrix metalloproteinase with substantial activity of decomposing enamel
 CC protein or similar activity. The products described in this invention can
 CC be used in the study of teeth development, enamel maturation process
 CC especially the formation of enamel or mineralisation of teeth matrix, and
 CC for medical, clinical and other applications. The proteins, their salts
 CC or peptide fragments, DNAs, promoters and inhibitors of biological
 CC activity of these proteins etc. can also be used as drugs.
 XX

SQ Sequence 483 AA;

Query Match 79.5%; Score 35; DB 20; Length 483;
 Best Local Similarity 100.0%; Pred. No. 53;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMYPTY 6
 :|||||
 Db 243 lmypty 248

RESULT 8
 AAW76253
 ID AAW76253 standard; Protein; 261 AA.

XX AC AAW76253;

XX DT 02-DEC-1998 (first entry)

XX DE Human matrilysin-like protein.

XX KW Matrilysin; inhibitor; tumour necrosis factor; screening; MMP; infection;
 KW matrix metalloproteinase; antagonist; diagnosis; underexpression; cancer;
 KW degeneration; extracellular matrix; arthritis; cardiovascular disease;
 KW cachexia; multiple sclerosis.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Peptide 1..22

FT FT /label= signal

FT Protein 23..261
 /label= matrilysin_like_protein

XX WO9831818-A2.

XX PN 23-JUL-1998.

XX 20-JAN-1998; 98WO-US00783.

XX 01-AUG-1997; 97US-0054541.

XX 21-JAN-1997; 97US-0034205.

XX 13-JUN-1997; 97US-0049607.

XX (HUMA-) HUMAN GENOME SCI INC.

XX PI Brewer L, Gentz R, Ni J, Rosen CA, Ruben SM;

XX WPI; 1998-414114/35.

XX N-PSDB; AAV61633.

XX Isolated nucleic acid encoding human metallo-protease(s) - used for
 PT diagnosis, treatment and prevention of, e.g. cancer, inflammation,
 PT neurological disease and infections

XX PS Claim 20a; Fig 4; 81pp; English.

XX This sequence represents a novel human matrilysin-like protein which is
 CC an inhibitor of the members of the matrix metalloproteinase, MMP.
 CC family. This protein can be used in assays to screen for agonists and
 CC antagonists and the nucleic acid is used as a probe for gene mapping, in

CC situ hybridisation and detection of corresponding genes in human tissue,
 CC and as sources of probes and primers for diagnosis. The protein and its
 CC antigenic fragments are used to raise antibodies (Ab) (which can be used
 CC for diagnosis in usual immunoassays or for in vivo imaging) and to screen
 CC for (ant)agonists. Agonists can be used to treat cancer or other
 CC conditions associated with underexpression of TNF-alpha. Antagonists are
 CC used to treat diseases associated with degeneration of the extracellular
 CC matrix (cancer, arthritis, cardiovascular disease, cachexia and multiple
 CC sclerosis).

XX SQ Sequence 261 AA;

Query Match 75.0%; Score 33; DB 19; Length 261;
 Best Local Similarity 83.3%; Pred. No. 68;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMYPTY 6
 :|||||
 Db 225 lmypty 230

RESULT 9
 AAB49983
 ID AAB49983 standard; Protein; 462 AA.

XX AC AAB49983;

XX DT 13-MAR-2001 (first entry)

XX DE Murine macrophage metalloelastase.

XX KW Mouse; macrophage metalloproteinase; elastin; matrix degrading enzyme;
 KW emphysema.

XX OS Mus sp.

XX PN US6150152-A.

XX PD 21-NOV-2000.

XX PF 28-MAY-1993; 93US-0068392.

XX PR 28-MAY-1993; 93US-0068392.

XX PA (UNIW) UNIV WASHINGTON.

XX PI Shapiro SD;

XX DR WPI; 2001-049090/06.

XX PT Novel human macrophage metalloelastase polynucleotides and polypeptides
 useful for measuring elastin degradation

XX PS Disclosure; Fig 1; 20pp; English.

XX The present invention provides the coding and protein sequences for the
 CC human macrophage metalloelastase. This is a matrix degrading
 CC metalloproteinase which has the ability to degrade elastin, and can be
 CC used to measure elastin degradation.

XX SQ Sequence 462 AA;

Query Match 75.0%; Score 33; DB 22; Length 462;
 Best Local Similarity 83.3%; Pred. No. 1.2e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMYPTY 6
 :|||||
 Db 228 lmypty 233

```

RESULT 10
AAW74618 standard; protein; 173 AA.
XX
AC AAW74618;
XX
DT 22-MAY-2001 (first entry)
XX
DE Human 72 kDa gelatinase catalytic domain protein sequence SEQ ID NO:8.
XX
DE Human; 92 kDa gelatinase; 72 kDa gelatinase; matrilysin; gelatinase;
XX
DE catalytic domain; 92 CD; 72 CD; matrilysin CD; elastin; vulnery;
XX
DE excess connective tissue removal; dermatological; keloid; scleroderma;
XX
DE post-operative fibrosis; intervertebral disc injection; fibrotic disease;
XX
DE hypertrophic scar; wound debridement; post-surgical adhesion;
XX
DE idiopathic pulmonary fibrosis.
XX
OS Homo sapiens.
XX
PI US6194189-B1.
XX
PD 27-FEB-2001.
XX
PF 16-DEC-1994; 94US-0357820.
XX
PR 16-DEC-1994; 94US-0357820.
XX
PA (UNIW ) UNIV WASHINGTON.
XX
PI Senior RM;
XX
WPI; 2001-243407/25.
XX
New gelatinase truncated mutant useful for treating disorders requiring
XX
the removal of excess connective tissue, e.g. keloids, post-operative
XX
fibrosis, intervertebral disc injections, hypertrophic scars -
XX
Example; Column 13-16; l1pp; English.
XX
The present invention describes a cDNA sequence which encodes the
XX
truncated mutant of the 92 kDa gelatinase having an amino acid sequence
XX
as given in AAW74618 consisting of residues 106-216 fused to residues
XX
391-443 of the parent molecule. The truncated mutant has dermatological
XX
and vulnery activities. The truncated mutant is useful for treating
XX
disorders requiring the removal of excess connective tissue,
XX
e.g., keloids, post-operative fibrosis, intervertebral disc injections,
XX
hypertrophic scars, wound debridement, post-surgical adhesions and
XX
various fibrotic diseases (including scleroderma, idiopathic pulmonary
XX
fibrosis). The truncated mutant is catalytically active compared to
XX
the full protein. Unlike the full protein, the truncated mutant is
XX
essentially inactive against insoluble elastin, and does not require
XX
activation to be enzymatically active. The present sequence represents
XX
a 72 kDa gelatinase catalytic domain (72 CD) containing protein, which
XX
is given in the exemplification of the present invention.
XX
SQ Sequence 173 AA;

Query Match 72.7%; Score 32; DB 22; Length 173;
Best Local Similarity 83.3%; Pred. No. 68;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LMVPTY 6
Db 137 vmypty 142

RESULT 11
AAW68682 standard; protein; 173 AA.
XX
AC AAW68682;
XX
DT 02-MAY-2001 (first entry)
XX
DE Human matrilysin protein.
XX
DE Matrilysin; inhibitor; tumour necrosis factor; screening; MMP; infection;
XX
DE matrix metalloproteinase; antagonist; diagnosis; underexpression; cancer;
XX
DE degeneration; extracellular matrix; arthritis; cardiovascular disease;
XX
DE cachexia; multiple sclerosis.
XX
OS Homo sapiens.
XX

```

```

DT 02-MAY-2001 (first entry)
XX
DE Matrilysin catalytic domain.
XX
DE Gelatinase; excess connective tissue removal; matrilysin;
XX
DE matrix metalloproteinase.
XX
OS Unidentified.
XX
PN US6184021-B1.
XX
PD 06-FEB-2001.
XX
PF 19-MAY-1995; 95US-0444628.
XX
PR 16-DEC-1994; 94US-0357820.
XX
PA (UNIW ) UNIV WASHINGTON.
XX
PI Senior RM;
XX
WPI; 2001-202001/20.
XX
New truncated mutant of 92 kDa gelatinase which is catalytically
XX
active, but is inactive against insoluble elastin, useful for treating
XX
disorders requiring the removal of excess connective tissues such as
XX
keloids -
XX
Examples; Fig 2; l1pp; English.
XX
The present invention relates to a truncated mutant (92 CD) of the 92
XX
kilo Dalton (kDa) gelatinase (see AAW68681). The truncated protein is
XX
useful for treating disorders requiring the removal of excess connective
XX
tissue, e.g. keloids, post-operative fibrosis, intervertebral disc
XX
injections, hypertrophic scars, wound debridement, post-surgical
XX
adhesions and various fibrotic diseases (scleroderma, idiopathic
XX
pulmonary fibrosis). Gelatinase is a matrix metalloproteinase and is also
XX
known as gelatinase B and MMP-9. The truncated protein is catalytically
XX
active comparable to the full protein but unlike the full protein is
XX
essentially inactive against insoluble elastin. The present sequence is
XX
the catalytic domain of matrilysin, which was used in a sequence homology
XX
alignment with the 92 CD protein of the present invention. Matrilysin is
XX
also a matrix metalloproteinase.
XX
SQ Sequence 173 AA;

Query Match 72.7%; Score 32; DB 22; Length 173;
Best Local Similarity 83.3%; Pred. No. 68;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LMVPTY 6
Db 137 vmypty 142

RESULT 12
AAW76254 standard; protein; 264 AA.
XX
ID AAW76254 standard; protein; 264 AA.
XX
AC AAW76254;
XX
DT 02-DEC-1998 (first entry)
XX
DE Human matrilysin protein.
XX
DE Matrilysin; inhibitor; tumour necrosis factor; screening; MMP; infection;
XX
DE matrix metalloproteinase; antagonist; diagnosis; underexpression; cancer;
XX
DE degeneration; extracellular matrix; arthritis; cardiovascular disease;
XX
DE cachexia; multiple sclerosis.
XX
OS Homo sapiens.
XX

```

PN W09831818-A2.
 XX
 PD 23-JUL-1998.
 XX
 XX PF
 XX 20-JAN-1998; 98WO-US00783.
 XX
 XX 01-AUG-1997; 97US-0054541.
 PR 21-JAN-1997; 97US-0034205.
 PR 13-JUN-1997; 97US-0049607.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Brewer L, Gentz R, Ni J, Rosen CA, Ruben SM;
 PI WPI; 1998-414114/35.
 XX
 XX Isolated nucleic acid encoding human metallo-protease(s) - used for
 PT diagnosis, treatment and prevention of, e.g. cancer, inflammation,
 PT neurological disease and infections
 PT
 XX Disclosure; Fig 5; 81pp; English.
 CC This sequence represents the human matrilysin protein which is an
 CC inhibitor of the members of the matrix metalloproteinase, MMP, family.
 CC This protein is used in the identification of a novel human
 CC matrilysin-like protein which can be used in assays to screen for
 CC agonists and antagonists and the nucleic acid is used as a probe for gene
 CC mapping, in situ hybridisation and detection of corresponding genes in
 CC human tissue, and as sources of probes and primers for diagnosis. The
 CC protein and its antigenic fragments are used to raise antibodies (Ab)
 CC (which can be used for diagnosis in usual immunoassays or for in vivo
 CC imaging) and to screen for (ant)agonists. Agonists can be used to treat
 CC cancer or other conditions associated with underexpression of TNF-alpha.
 CC Antagonists are used to treat diseases associated with degeneration of
 CC the extracellular matrix (cancer, arthritis, cardiovascular disease,
 CC cachexia and multiple sclerosis).
 XX
 XX Sequence 264 AA;

Query Match 72.7%; Score 32; DB 19; Length 264;
 Best Local Similarity 83.3%; Pred. No. 1.1e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LMYPYTY 6
 :|||||
 Db 231 vmypty 236

ULT 13
 AAR96211
 ID AAR96211 standard; Protein; 271 AA.

XX AAR96211;
 AC
 DT 20-NOV-1996 (first entry)
 DE Human recombinant matrilysin.
 XX
 XX Matrilysin; cell dispersal; enzyme; recombinant; lysis;
 KW cell separation.
 KW
 XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FH Peptide 1..21
 FT /label= sig_peptide
 FT 22..271
 FT /label= mat_protein
 FT
 XX JP08098693-A.
 PN 16-APR-1996.
 PD

XX 30-SEP-1994; 94JP-0259576.
 PF
 XX 30-SEP-1994; 94JP-0259576.
 PR
 XX (ORIY) ORIENTAL YEAST CO LTD.
 PA
 XX Kihira Y;
 PI
 XX WPI; 1996-245872/25.
 DR N-PSDB; AAT27538.
 DR
 XX Recombinant prodn. of precursor type human matrilysin - used for
 PT dispersing cells, e.g. for removing cells from the walls of culture
 PT vessels etc.
 PT
 XX Claim 7; Page 4; 7pp; Japanese.
 PS
 XX AAT27538 is a copy of the human matrilysin gene which can be
 CC engineered so it can be expressed in E. coli as an immature protein
 CC with a signal peptide directing secretion of the protein into the
 CC periplasm. The protein was shown to be expressed easily in E. coli in
 CC its active form. Matrilysin produced can be used as a cell-dispersal
 CC reagent e.g. to remove animal cells from the wall of a vessel such as
 CC a test tube, or to separate specific cells from various animal
 CC tissues to aid in the study of medical chemistry, biochemistry or
 CC pharmacy.
 CC
 XX Sequence 271 AA;
 SQ
 Query Match 72.7%; Score 32; DB 17; Length 271;
 Best Local Similarity 83.3%; Pred. No. 1.1e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LMYPYTY 6
 :|||||
 Db 235 vmypty 240

RESULT 14
 AAB43585
 ID AAB43585 standard; Protein; 297 AA.
 XX
 XX AAB43585;
 AC
 XX 08-FEB-2001 (first entry)
 DT
 XX Human cancer associated protein sequence SEQ ID NO:1030.
 DE
 XX Human; cancer associated gene; cancer antigen; detection; cancer;
 KW diagnosis; cytostatic; proliferative; vulnery; immunomodulator;
 KW antidiabetic; antisthmatic; antirheumatic; antithratic; antiviral;
 KW antiinflammatory; antithyroid; antiallergic; antibacterial; cardiac;
 KW dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
 KW vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
 KW immune disorder; hematopoietic cell disorder; autoimmune disorder;
 KW allergic reaction; graft versus host disease; organ rejection;
 KW haemostatic; thrombolytic; cardiovascular disorder; infection;
 KW neurological disease; drug screening.
 XX
 XX Homo sapiens.
 OS
 XX W0200055350-A1.
 PN
 XX 21-SEP-2000.
 PD
 XX 08-MAR-2000; 2000WO-US05882.
 PF
 XX 12-MAR-1999; 99US-0124270.
 PR
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX

PI Rosen CA, Ruben SM;
 XX WPI; 2000-587533/55.
 DR N-PSDB; AAC77794.
 XX
 PT Novel isolated nucleic acids comprising sequences encoding peptides
 PT useful for treating or diagnosing e.g. cancer.
 XX
 PS Claim 11; Page 1615-1616; 2352pp; English.
 XX
 CC AAC77607 to AAC78448 encode the human cancer associated proteins given
 CC in AAB43398 to AAB44239. The proteins can have activities based on the
 CC tissues and cells the genes are expressed in. Example of activities
 CC include: cytostatic; proliferative; vulnerable; immunomodulator;
 CC antidiabetic; antilasthmatic; antirheumatic; antiarthritic;
 CC antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
 CC dermatological; neuroprotective; cardiac; thrombolytic; coagulant;
 CC nontropic; vasotropic; antipsoriatic and antiangiogenic. The
 CC polynucleotides and polypeptides can be used for preventing, treating or
 CC ameliorating medical conditions and diagnosing pathological conditions.
 CC polynucleotides, polypeptides, antibodies, agonists and antagonists from
 CC the present invention may be used to treat immune disorders by activating
 CC inhibiting the proliferation, differentiation or mobilisation of
 CC immune cells, to treat disorders of haematopoietic cells, autoimmune
 CC disorders, allergic reactions, graft versus host disease and organ
 CC rejection, modulate haemostatic or thrombolytic activity, modulate
 CC bacterial or viral infections. The peptides, nucleotides, antibodies,
 CC agonists and antagonists may be also be used in drug screens. AAC78449 to
 CC AAC78457 and AAB44240 represent sequences used in the exemplification of
 CC the present invention.
 XX
 SQ Sequence 297 AA;

Query Match 72.7%; Score 32; DB 21; Length 297;
 Best Local Similarity 83.3%; Pred. No. 1.2e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMVPTY 6
 :|||||
 Db 261 vmypty 266

RESULT 15
 AAR87016
 ID AAR87016 standard; Protein; 370 AA.
 AAR87016;
 19-MAR-1996 (first entry)
 DE Type I matrix metalloprotease.
 XX
 KW Type I matrix metalloprotease; MMP-I; collagenase; nematode;
 KW disease resistance; transgenic plant; crop improvement;
 KW nematocide; biological control; potato; Solanum tuberosum.
 XX
 OS Not specified.
 XX
 PN W09530017-A1.
 XX
 PD 09-NOV-1995.
 XX
 PF 26-APR-1995; 95WO-GB00944.
 XX
 PR 29-APR-1994; 94EP-0303168.
 XX
 PA (UNIL) UNILEVER NV.
 XX (UNIL) UNILEVER PLC.
 XX
 PI Shields R, Stratford R;
 XX

DR WPI; 1995-393088/50.
 DR N-PSDB; AAT07301.
 XX
 PT Nucleic acid for increasing plant resistance to nematodes - encodes
 PT a polypeptide with collagenase activity.
 XX
 PS Example; Page 23-24; 41pp; English.
 XX
 CC The mature form of type I matrix metalloprotease (MMP-I) (AAR87016)
 CC has collagenase activity that degrades native collagens types I,
 CC II and III. DNA (AAT07301) coding for mature MMP-I was modified
 CC for expression in plants by addition of a start codon and a patatin
 CC leader sequence. Transgenic potato cv. Desiree and Maris Piper
 CC plants were obtd. that showed high levels of resistance to the
 CC nematode Globodera pallida.
 XX
 SQ Sequence 370 AA;

Query Match 72.7%; Score 32; DB 16; Length 370;
 Best Local Similarity 71.4%; Pred. No. 1.5e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LMVPTYL 7
 :|||||
 Db 136 lmyptyl 142

Search completed: September 4, 2001, 15:50:25
 Job time: 163 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2001, 15:47:57 ; Search time 40.98 Seconds
(without alignments)
4.020 Million cell updates/sec

Title: US-09-630-345-1
Perfect score: 44
Sequence: 1 LMYPTYLK 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents:AA:*
1: /cgn2_6/ptodata/2/iaa/5A-COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B-COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A-COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B-COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS-COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	100.0	8	1 US-07-620-410-3	Sequence 3, Appl1
2	33	75.0	261	3 US-09-009-156-5	Sequence 5, Appl1
3	33	75.0	462	4 US-08-068-392-3	Sequence 3, Appl1
4	33	75.0	462	4 US-08-396-988-3	Sequence 3, Appl1
5	32	72.7	173	4 US-08-444-628-8	Sequence 8, Appl1
6	32	72.7	173	4 US-08-357-820-8	Sequence 8, Appl1
7	32	72.7	264	3 US-09-009-156-6	Sequence 6, Appl1
8	32	72.7	267	4 US-08-448-489-18	Sequence 18, Appl1
9	32	72.7	271	3 US-08-896-062-2	Sequence 2, Appl1
10	31	70.5	8	2 US-08-922-267A-65	Sequence 65, Appl1
11	31	70.5	231	4 US-08-448-489-19	Sequence 19, Appl1
12	31	70.5	459	1 US-08-630-592-2	Sequence 2, Appl1
13	31	70.5	459	1 US-08-714-991-2	Sequence 2, Appl1
14	31	70.5	459	3 US-09-032-365A-2	Sequence 7, Appl1
15	31	70.5	460	1 US-08-630-592-7	Sequence 7, Appl1
16	31	70.5	460	1 US-08-714-991-7	Sequence 7, Appl1
17	31	70.5	460	3 US-09-032-365A-8	Sequence 8, Appl1
18	31	70.5	469	3 US-08-704-711A-16	Sequence 16, Appl1
19	31	70.5	469	4 US-08-448-489-12	Sequence 12, Appl1
20	31	70.5	505	1 US-08-631-200-2	Sequence 2, Appl1
21	31	70.5	505	1 US-08-630-592-4	Sequence 4, Appl1
22	31	70.5	505	1 US-08-714-991-4	Sequence 4, Appl1
23	31	70.5	505	2 US-08-829-553-2	Sequence 2, Appl1
24	31	70.5	505	2 US-08-922-267A-2	Sequence 2, Appl1
25	31	70.5	505	2 US-08-936-707A-2	Sequence 2, Appl1
26	31	70.5	505	2 US-08-936-706A-2	Sequence 2, Appl1
27	31	70.5	505	3 US-09-248-203-2	Sequence 2, Appl1

28	31	70.5	505	3 US-09-032-365A-4	Sequence 4, Appl1
29	31	70.5	505	4 US-08-812-824-3	Sequence 3, Appl1
30	31	70.5	505	4 US-09-406-071-2	Sequence 2, Appl1
31	31	70.5	506	1 US-08-631-200-8	Sequence 8, Appl1
32	31	70.5	506	2 US-08-829-553-8	Sequence 8, Appl1
33	31	70.5	506	2 US-08-922-267A-8	Sequence 8, Appl1
34	31	70.5	506	2 US-08-936-707A-8	Sequence 8, Appl1
35	31	70.5	506	2 US-08-936-706A-8	Sequence 8, Appl1
36	31	70.5	506	3 US-09-248-203-8	Sequence 8, Appl1
37	31	70.5	506	3 US-09-032-365A-62	Sequence 62, Appl1
38	31	70.5	506	4 US-08-812-824-4	Sequence 4, Appl1
39	31	70.5	506	4 US-09-406-071-8	Sequence 8, Appl1
40	31	70.5	512	3 US-09-032-365A-60	Sequence 60, Appl1
41	31	70.5	518	3 US-09-032-365A-58	Sequence 58, Appl1
42	31	70.5	561	1 US-08-714-991-27	Sequence 27, Appl1
43	31	70.5	561	3 US-09-032-365A-10	Sequence 10, Appl1
44	30	68.2	73	1 US-08-379-538-7	Sequence 7, Appl1
45	30	68.2	444	1 US-09-178-002-2	Sequence 2, Appl1

ALIGNMENTS

RESULT 1
US-07-620-410-3
; Sequence 3, Application US/07620410
; Patent No. 5217953
; GENERAL INFORMATION:
; APPLICANT: Gozes, Iillana
; APPLICANT: Brennenman, Douglas E.
; APPLICANT: Fridkin, Matti
; APPLICANT: Moody, Terry
; TITLE OF INVENTION: A NOVEL VASOACTIVE INTESTINAL PEPTIDE
; TITLE OF INVENTION: ANTAGONIST
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
; STREET: 1615 L. Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/620,410
; FILING DATE: 19901130
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Scott, Watson T.
; REGISTRATION NUMBER: 26,581
; REFERENCE/DOCKET NUMBER: WTS/5683/82679
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 248453 CUSH
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
US-07-620-410-3

Query Match 100.0%; Score 44; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LMYPTYLK 8
|||||||

Db 1 LMYPTLK 8

RESULT 2
US-09-009-156-5
; Sequence 5, Application US/09009156
; Patent No. 6046031
; GENERAL INFORMATION:
; APPLICANT: Ni, Jian
; APPLICANT: Ruben, Steven
; APPLICANT: Brewer, Laurie
; APPLICANT: Gentz, Elaine
; TITLE OF INVENTION: No. 6046031lel Metalloproteinases
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: human Genome Sciences, Inc
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: US
; ZIP: 20850

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/009,156
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/034,205
; FILING DATE: 21-JAN-1997
; PRIOR APPLICATION DATA: US 60/049,607
; FILING DATE: 13-JUN-1997
; PRIOR APPLICATION DATA: US 60/054,541
; FILING DATE: 01-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PF376
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-309-8504
; TELEFAX: 301-309-8439
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 261 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-009-156-5

Query Match 75.0%; Score 33; DB 3; Length 261;
Best Local Similarity 83.3%; Pred. No. 38;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LMYPTK 6

Db 225 LMYPTK 230

RESULT 3
US-08-068-392-3
; Sequence 3, Application US/08068392
; Patent No. 6150152
; GENERAL INFORMATION:
; APPLICANT: Shapiro, Steven M.
; TITLE OF INVENTION: Human Macrophage Metalloproteinase
; NUMBER OF SEQUENCES: 3

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scott J. Meyer, Monsanto Co., A3SM
; STREET: 800 N. Lindbergh Blvd.
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63167
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/068,392
; FILING DATE: 19930528
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyer, Scott J.
; REGISTRATION NUMBER: 25275
; REFERENCE/DOCKET NUMBER: 07-24(12406)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)694-3117
; TELEFAX: (314)694-5435
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 462 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-068-392-3

Query Match 75.0%; Score 33; DB 4; Length 462;
Best Local Similarity 83.3%; Pred. No. 69;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LMYPTK 6

Db 228 LMYPTK 233

RESULT 4
US-08-396-988-3
; Sequence 3, Application US/08396988
; Patent No. 6204043
; GENERAL INFORMATION:
; APPLICANT: Shapiro, Steven M.
; TITLE OF INVENTION: Human Macrophage Metalloproteinase
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scott J. Meyer, Monsanto Co., A3SM
; STREET: 800 N. Lindbergh Blvd.
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63167
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/396,988
; FILING DATE: 01-MAR-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/068,392
; FILING DATE: 28-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyer, Scott J.
; REGISTRATION NUMBER: 25275
; REFERENCE/DOCKET NUMBER: 07-24(12406)A
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (314)694-3117
TELEFAX: (314)694-5435
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 462 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-396-988-3

Query Match 75.0%; Score 33; DB 4; Length 462;
Best Local Similarity 83.3%; Pred. No. 69;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMPTY 6
:|||||
Db 228 IMPTY 233

US-08-444-628-8
Sequence 8, Application US/08444628
Patent No. 6184021
GENERAL INFORMATION:
APPLICANT: Senior, Robert M.
TITLE OF INVENTION: Catalytically-active Gelatinase Mutant
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scott J. Meyer, Monsanto/Searle, A3SG
STREET: 800 N. Lindbergh Blvd.
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63167
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/444,628
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/357,820
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Meyer, Scott J.
REGISTRATION NUMBER: 25,275
REFERENCE/DOCKET NUMBER: WU-2855
TELEPHONE: (314)694-3117
TELEFAX: (314)694-5435
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 173 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-444-628-8

Query Match 72.7%; Score 32; DB 4; Length 173;
Best Local Similarity 83.3%; Pred. No. 38;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMPTY 6
:|||||
Db 137 VMPTY 142

RESULT 6
US-08-357-820-8
Sequence 8, Application US/08357820
Patent No. 6194189
GENERAL INFORMATION:
APPLICANT: Senior, Robert M.
TITLE OF INVENTION: Catalytically-active Gelatinase Mutant
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scott J. Meyer, Monsanto/Searle, A3SG
STREET: 800 N. Lindbergh Blvd.
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63167
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/357,820
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Meyer, Scott J.
REGISTRATION NUMBER: 25,275
REFERENCE/DOCKET NUMBER: WU-2855
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314)694-3117
TELEFAX: (314)694-5435
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 173 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-357-820-8

Query Match 72.7%; Score 32; DB 4; Length 173;
Best Local Similarity 83.3%; Pred. No. 38;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMPTY 6
:|||||
Db 137 VMPTY 142

RESULT 7
US-09-009-156-6
Sequence 6, Application US/09009156
Patent No. 6046031
GENERAL INFORMATION:
APPLICANT: Ni, Jian
APPLICANT: Ruben, Steven
APPLICANT: Brewer, Laurie
APPLICANT: Gentz, Reiner
TITLE OF INVENTION: No. 6046031el Metalloproteinases
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: US
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

```

; CURRENT APPLICATION DATA:
; FILING DATE: US/09/009,156
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; FILING DATE: 21-JAN-1997
; PRIOR APPLICATION NUMBER: US 60/034,205
; FILING DATE: 13-JUN-1997
; PRIOR APPLICATION NUMBER: US 60/049,607
; FILING DATE: 01-AUG-1997
; PRIOR APPLICATION DATA: US 60/054,541
; FILING DATE: 01-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Andrews
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PF376
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-309-8504
; TELEFAX: 301-309-8439
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 264 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-009-156-6

```

```

Query Match 72.7%; Score 32; DB 3; Length 264;
Best Local Similarity 83.3%; Pred. No. 59;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 LMPTY 6
   :|||||
Db 231 VMPTY 236

```

```

RESULT 8
US-08-448-489-18
; Sequence 18, Application US/08448489
; Patent No. 6184022
; GENERAL INFORMATION:
; APPLICANT: SEIKI, Motoharu
; APPLICANT: SATO, Hiroshi
; APPLICANT: SHINAGAWA, Akira
; TITLE OF INVENTION: NOVEL METALLOPROTEINASE AND ENCODING DNA THEREFOR
; FILE REFERENCE: 55-290P
; CURRENT APPLICATION NUMBER: US/08/448,489
; CURRENT FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Known Member of
; OTHER INFORMATION: Matrix Metalloproteinase Family
US-08-448-489-18

```

```

Query Match 72.7%; Score 32; DB 4; Length 267;
Best Local Similarity 83.3%; Pred. No. 60;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 LMPTY 6
   :|||||
Db 231 VMPTY 236

```

```

RESULT 9

```

```

US-08-896-062-2
; Sequence 2, Application US/08896062
; Patent No. 6010893
; GENERAL INFORMATION:
; APPLICANT: KIHIRA, Yasunori
; TITLE OF INVENTION: PROCESS
; TITLE OF INVENTION: BY MEANS OF RECOMBINANT DNA
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Browdy and Neimark
; STREET: 419 Seventh Street N.W., Ste. 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States of America
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/896,062
; FILING DATE: 17-JUL-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/530,984
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Browdy, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: KIHIRA-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-5197
; TELEFAX: (202) 737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 271 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-896-062-2

```

```

Query Match 72.7%; Score 32; DB 3; Length 271;
Best Local Similarity 83.3%; Pred. No. 61;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 LMPTY 6
   :|||||
Db 235 VMPTY 240

```

```

RESULT 10
US-08-922-267A-65
; Sequence 65, Application US/08922267A
; Patent No. 5861239
; GENERAL INFORMATION:
; APPLICANT: Kieyn, Patrick W.
; APPLICANT: Moore, Karen J.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; NUMBER OF SEQUENCES: 82
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/922,267A
FILING DATE: 2-SEP-1997
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/829,553
FILING DATE: 28-MAR-1997
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/631,200
FILING DATE: 12-APR-1996
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-085
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 8 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-922-267A-65

Query Match 70.5%; Score 31; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
| | | | |
DB 2 MYPTY 6

RESULT 11
US-08-448-489-19
Sequence 19, Application US/08448489
Patent No. 6184022
GENERAL INFORMATION:
APPLICANT: SEIKI, Motoharu
APPLICANT: SATO, Hiroshi
APPLICANT: SHINAGAWA, Akira
TITLE OF INVENTION: NOVEL METALLOPROTEINASE AND ENCODING DNA THEREFOR
FILE REFERENCE: 55-290P
CURRENT APPLICATION NUMBER: US/08/448,489
CURRENT FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 19
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 19
LENGTH: 231
TYPE: PRT
ORGANISM: Unknown
FEATURE:
OTHER INFORMATION: Description of Unknown Organism: Known Member of
OTHER INFORMATION: Matrix Metalloproteinase Family
US-08-448-489-19

Query Match 70.5%; Score 31; DB 4; Length 231;
Best Local Similarity 83.3%; Pred. NO. 80;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMYPYTF 6
| | | | |
DB 138 LMYPYTF 143

RESULT 12
US-08-630-592-2
Sequence 2, Application US/08630592
Patent No. 5770432
GENERAL INFORMATION:
APPLICANT: Nishina, Patsy
APPLICANT: No. 5770432enTrauth, Konrad
APPLICANT: Naggett, Juergen
APPLICANT: No. 5770432th, Michael
TITLE OF INVENTION: Obesity Associated Genes
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
STREET: 3400 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 941114187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PCDOS/MSDOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/630,592
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sherwood, Pamela J.
REGISTRATION NUMBER: 36,677
REFERENCE/DOCKET NUMBER: A59504/BIR/PJS
TELEPHONE: (415) 7811989
TELEFAX: (415) 3983249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 459 amino acids
TYPE: amino acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-630-592-2

Query Match 70.5%; Score 31; DB 1; Length 459;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
| | | | |
DB 233 MYPTY 237

RESULT 13
US-08-714-991-2
Sequence 2, Application US/08714991
Patent No. 5776762
GENERAL INFORMATION:
APPLICANT: NORTH, Michael
APPLICANT: NISHINA, Patsy
APPLICANT: No. 5776762en-Trauth, Konrad
APPLICANT: NAGGETT, Juergen
TITLE OF INVENTION: OBESITY ASSOCIATED GENES
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/714,991
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: SHERWOOD, Pamela J.
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: A-59504-1/PJS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-494-8770
; TELEFAX: 415-494-8771
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 459 amino acids
; TYPE: amino acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-714-991-2

```

```

Query Match 70.5%; Score 31; DB 1; Length 459;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 2 MYPTY 6
Db 233 MYPTY 237

```

```

RESULT 14
US-09-630-345A-2
; Sequence 2, Application US/09032365A
; Patent No. 6114502
; GENERAL INFORMATION:
; APPLICANT: Nishina, Patsy
; APPLICANT: Nishina, Patsy
; APPLICANT: Nagger, Juergen
; APPLICANT: Nagger, Juergen
; TITLE OF INVENTION: GENE FAMILY ASSOCIATED WITH
; TITLE OF INVENTION: NEUROSENSORY DEFECTS
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bozicevic & Reed, LLP
; STREET: 285 Hamilton Avenue, Suite 200
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/032,365A
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J.
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: SEQ-2CIP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-327-3400

```

```

; TELEFAX: 650 327-3231
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 459 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-032-365A-2

```

```

Query Match 70.5%; Score 31; DB 3; Length 459;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 2 MYPTY 6
Db 233 MYPTY 237

```

```

RESULT 15
US-08-630-592-7
; Sequence 7, Application US/08630592
; Patent No. 5770432
; GENERAL INFORMATION:
; APPLICANT: Nishina, Patsy
; APPLICANT: Nishina, Patsy
; APPLICANT: Nagger, Juergen
; APPLICANT: Nagger, Juergen
; TITLE OF INVENTION: Obesity Associated Genes
; TITLE OF INVENTION: Obesity Associated Genes
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ELLEN, HOBBACH, TEST, ALBRITTON & HERBERT
; STREET: 3400 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94114187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PCDOS/MSDOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/630,592
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J.
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: A59504/BIK/PJS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 7811989
; TELEFAX: (415) 3983249
; TELEFAX: 910 277299
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 460 amino acids
; TYPE: amino acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-630-592-7

```

```

Query Match 70.5%; Score 31; DB 1; Length 460;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 2 MYPTY 6
Db 234 MYPTY 238

```

Search completed: September 4, 2001, 15:51:12
Job time: 195 sec

THIS PAGE BLANK (user to)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2001, 15:48:17 ; Search time 49.08 Seconds
(without alignments)
12.416 Million cell updates/sec

Title: US-09-630-345-1
Perfect score: 44
Sequence: 1 LMPTYLK 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

al number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41	93.2	384	2	I51267 collagenase (EC 3.4.24.-) - bullfrog
2	36	81.8	218	2	E86750 hypothetical prote
3	36	81.8	340	1	MMBE5 cell fusion protei
4	36	81.8	393	2	B86644 transporter ybfD f
5	36	81.8	436	2	C69764 4-aminobutyrate am
6	35	79.5	378	2	F86307 hypothetical prote
7	35	79.5	483	2	JC5743 matrix metalloprot
8	34	77.3	269	2	D69011 hypothetical prote
9	33	75.0	262	2	C81384 shikimate 5-dehydr
10	33	75.0	462	2	A42401 macrophage elastas
11	33	75.0	511	2	T19496 hypothetical prote
12	32	72.7	187	2	S69693 hypothetical prote
13	32	72.7	267	1	KCHUM matrilysin (EC 3.4
14	32	72.7	267	2	A57490 matrilysin (EC 3.4
15	32	72.7	349	2	JU0458 peroxidase (EC 1.1
16	32	72.7	352	2	T21084 hypothetical prote
17	32	72.7	425	2	T32537 hypothetical prote
18	32	72.7	468	1	KCRBI interstitial colla
19	32	72.7	469	1	KCPGI interstitial colla
20	32	72.7	593	2	JC4884 organic cation tra
21	32	72.7	1263	2	F44466 DNA-directed RNA p
22	31	70.5	247	2	B85515 hypothetical prote
23	31	70.5	262	2	T29698 hypothetical prote
24	31	70.5	291	2	B86356 hypothetical prote
25	31	70.5	301	2	S57923 SEC14 protein - ye
26	31	70.5	302	2	H96811 protein F3F9.20 (i
27	31	70.5	385	2	E86359 Similar to seed ma
28	31	70.5	402	2	S47329 OXAL protein precu
29	31	70.5	469	1	KCHUI interstitial colla

30	31	70.5	469	1	KCB0I interstitial colla
31	31	70.5	478	2	A86677 phospho-beta-gluc
32	31	70.5	495	1	PIWLB L1 protein - bovin
33	31	70.5	501	1	PIWLB2 L1 protein - bovin
34	31	70.5	501	1	PIWLEP L1 protein - Europ
35	31	70.5	502	2	T19708 hypothetical prote
36	31	70.5	503	2	T29175 hypothetical prote
37	31	70.5	505	2	S68518 tub protein, brain
38	31	70.5	513	1	PIWLDP L1 protein - deer
39	31	70.5	516	2	T15633 hypothetical prote
40	31	70.5	565	2	H69363 hypothetical prote
41	31	70.5	581	2	T16915 conserved hypothet
42	31	70.5	625	2	F70459 kpsc protein - Esc
43	31	70.5	662	2	S42826 hypothetical prote
44	31	70.5	675	2	D48492
45	31	70.5	1068	2	S01519

ALIGNMENTS

RESULT 1
I51267
collagenase (EC 3.4.24.-) - bullfrog
C:Species: Rana catesbeiana (bullfrog)
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 22-Jun-1999
C:Accession: I51267
R:Oofusa, K.; Yomori, S.; Yoshizato, K.
Int. J. Dev. Biol. 38, 345-350, 1994
A:Title: Regionally and hormonally regulated expression of genes of collagen and coll
A:Reference number: I51267; MUID:95071832
A:Accession: I51267
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-384 <OOF>
A:Cross-references: GB:S75623; NID:g913070; PIDN:AAB32661.1; PID:g913071
C:Superfamily: Interstitial collagenase; hemoxen repeat homology; matrix metallopro
F:53-231/Domain: matrix metalloproteinase; zinc; zymogen
F:236-381/Domain: hemoxen repeat homology <PXM>
F:81,189,193,199/Binding site: zinc, catalytic (Cys, His, His, His) (inhibited) #stat
F:189,193,199/Binding site: zinc, catalytic (His) (active) #status predicted
F:190/Active site: Glu #status predicted

Query Match 93.2%; Score 41; DB 2; Length 384;
Best Local Similarity 87.5%; Pred. No. 1.3;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LMPTYLK 8
DB 206 LMPTYLK 213

RESULT 2
E86750
hypothetical protein ykdb [imported] - Lactococcus lactis subsp. lactis (strain IL140
C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 23-Mar-2001
C:Accession: E86750
R:Botolin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissensbach, J.; Eh
Genome Res. in press, 2001
A:Title: The complete genome sequence of the lactic acid bacterium.
A:Reference number: A86625
A:Accession: E86750
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-218 <STO>
A:Cross-references: GB:AE005176; NID:g12723952; PIDN:AAK05103.1; GSPDB:GN00146
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: ykdb

Query Match 81.8%; Score 36; DB 2; Length 218;
 Best Local Similarity 75.0%; Pred. No. 7.4;
 Matches 6; Conservative 1; Mismatches 0; Gaps 0;

Qy 1 LMYPYTLK 8
 | |||||
 Db 202 LTYPTYLK 209

RESULT 3

MMB5
 C:cell fusion protein precursor - human herpesvirus 3
 C:Species: human herpesvirus 3, varicella-zoster virus
 C:Date: 30-Sep-1998 #sequence_revision 30-Sep-1998 #text_change 16-Jul-1999
 C:Accession: E27212
 R:Davidson, A.J.; Scott, J.B.
 J. Gen. Virol. 67, 1759-1816, 1986
 A:Title: The complete DNA sequence of varicella-zoster virus.
 A:Reference number: A27345; MUID:86306657

A:Accession: E27212
 A:Molecule type: DNA
 A:Residues: 1-340 <DAV>
 A:Cross-references: EMBL:X04370; NID:G59989; PIDN:CAA27888.1; PID:G59994
 C:Genetics:
 A:Gene: 5

C:Superfamily: herpesvirus cell fusion protein
 C:Keywords: membrane fusion; transmembrane protein
 F:1-31/Domain: signal sequence #status predicted <SIG>
 F:32-340/Product: cell fusion protein #status predicted <CFP>
 F:115-137/Domain: transmembrane #status predicted <TM1>
 F:220-238/Domain: transmembrane #status predicted <TM2>
 F:251-269/Domain: transmembrane #status predicted <TM3>
 F:307-322/Domain: transmembrane #status predicted <TM5>

Query Match 81.8%; Score 36; DB 1; Length 340;
 Best Local Similarity 62.5%; Pred. No. 12;
 Matches 5; Conservative 3; Mismatches 0; Indels 0;

Qy 1 LMYPYTLK 8
 | |||||
 Db 243 LLYPTYIR 250

RESULT 4

B86644
 C:transporter ybfD [imported] - Lactococcus lactis subsp. lactis (strain ILL1403)
 C:Species: Lactococcus lactis subsp. lactis
 C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 23-Mar-2001
 C:Accession: B86644
 R:Belotin, A.; Wincker, P.; Mager, S.; Jaillon, O.; Malarre, K.; Weissenbach, J.; Ehrlich
 Genome Res. in press, 2001
 A:Title: The complete genome sequence of the lactic acid bacterium.

A:Reference number: B86625
 A:Accession: B86644
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-393 <STO>
 A:Cross-references: GB:AE005176; NID:912723004; PIDN:AAK04252.1; GSPDB:GN00146
 A:Experimental source: strain ILL1403
 C:Genetics:
 A:Gene: ybfD

Query Match 81.8%; Score 36; DB 2; Length 393;
 Best Local Similarity 85.7%; Pred. No. 14;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LMYPYTL 7
 | |||||
 Db 277 VMVPTYL 283

RESULT 5

C69764
 C:aminobutyrate aminotransferase homolog ycnG - Bacillus subtilis
 C:Species: Bacillus subtilis
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
 C:Accession: C69764
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber
 C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
 A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,
 Nature 390, 249-256, 1997
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gal
 lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M
 Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino
 A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau
 Y., M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete
 Rieger, M.; Rivoita, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, F.; Sato, T.; Scanl
 A:Authors: Schleich, S.; Schroeter, T.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Se
 keuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiya
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida
 A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis
 A:Reference number: A69580; MUID:98044033
 A:Accession: C69764
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-436 <KUN>

A:Cross-references: GB:299106; GB:AL009126; NID:G2632653; PIDN:CAB12198.1; PID:G26326
 A:Experimental source: strain 168
 C:Genetics:
 A:Gene: ycnG
 C:Superfamily: ornithine--oxo-acid aminotransferase

Query Match 81.8%; Score 36; DB 2; Length 436;
 Best Local Similarity 62.5%; Pred. No. 15;
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LMYPYTLK 8
 | |||||
 Db 87 MMYPTYIE 94

RESULT 6

F86307
 C:hypothetical protein AAD50017.1 [Imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: F86307
 R:Rheologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Luros, J.S.; Maitl, R.; Marzla
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719
 A:Accession: F86307
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-378 <STO>
 A:Cross-references: GB:AE005172; NID:G5734752; PIDN:AAD50017.1; GSPDB:GN00141
 C:Genetics:
 A:Map position: 1

Query Match 79.5%; Score 35; DB 2; Length 378;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMVPTY 6
| | | | |
Db 117 LMVPTY 122

RESULT 7
JC5743
matrix metalloproteinase (EC 3.4.24.-) precursor - pig
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 09-Dec-1997 #sequence_revision 23-Jan-1998 #text_change 22-Jun-1999
C:Accession: JC5743
R:Barlett, J.D.; Simmer, J.P.; Xue, J.; Margolis, H.C.; Moreno, E.C.
Gene 183, 123-126, 1996
A:Title: Molecular cloning and mRNA tissue distribution of a novel matrix metalloproteinase
A:Reference number: JC5743; MUID:97149288
A:Accession: JC5743
A:Molecule type: mRNA
A:Residues: 1-483 <BAR>
A:Cross-references: GB:U54825; NID:gl800212; PIDN:AAB41396.1; PID:gl800213
A:Experimental source: enamel organ
A:Comment: This enzyme plays a role in enamel biomineralization and development.
A:Superfamily: Interstitial collagenase; hemopexin repeat homology: matrix metalloproteinase
C:Keywords: hydrolase; metalloproteinase; zinc
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-483/Product: matrix metalloproteinase #status predicted <WAT>
F:68-271/Domain: matrix metalloproteinase homology <WMP>
F:290-483/Domain: hemopexin repeat homology <PXN>
F:100,226,230,236/Binding site: zinc, catalytic (Cys, His, His) (inhibited) #status
F:226,230,236/Binding site: zinc, catalytic (His) #status predicted
F:227/Active site: Glu #status predicted

Query Match 79.5%; Score 35; DB 2; Length 483;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMVPTY 6
| | | | |
Db 243 LMVPTY 248

RESULT 8
D69011
hypothetical protein MTH1087 - Methanobacterium thermoautotrophicum (strain Delta H)
C:Species: Methanobacterium thermoautotrophicum
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C:Accession: D69011
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
Glu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.
S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: func
A:Reference number: A69000; MUID:98037514
A:Accession: D69011
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-269 <MTH>
A:Cross-references: GB:AE000879; GB:AE000666; NID:g2622175; PIDN:AAB85576.1; PID:g262218
A:Experimental source: strain Delta H
C:Genetics:
A:Gene: MTH1087

Query Match 77.3%; Score 34; DB 2; Length 269;
Best Local Similarity 62.5%; Pred. No. 23;
Matches 5; Conservative 2; Mismatches 0; Indels 1; Gaps 0;

QY 1 LMVPTYLK 8
: | | | | :
Db 173 MTPPTYIK 180

RESULT 9

C81384
shikimate 5-dehydrogenase (EC 1.1.1.25) Cj0405 [imported] - Campylobacter jejuni (str
C:Species: Campylobacter jejuni
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 26-May-2000
C:Accession: C81384
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chil
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Ba
Nature 403, 665-668, 2000
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals
A:Reference number: A81250; MUID:20150912
A:Accession: C81384
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-262 <PAR>
A:Cross-references: GB:AL139075; GB:AL111168; NID:g6967817; PIDN:CAB74241.1; PID:g696
A:Experimental source: serotype O2, strain NCTC 11168
C:Genetics:
A:Gene: aroE; Cj0405
A:Superfamily: shikimate dehydrogenase; shikimate dehydrogenase homology
C:Keywords: oxidoreductase

Query Match 75.0%; Score 33; DB 2; Length 262;
Best Local Similarity 85.7%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LMVPTYL 7
| | | | |
Db 153 LAYPTVL 159

RESULT 10
A42401
macrophage elastase (EC 3.4.24.-) precursor - mouse
N:Alternate names: matrix metalloproteinase 12 (MMP12)
C:Species: Mus musculus (house mouse)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
C:Accession: A42401
R:Shapiro, S.D.; Griffin, G.L.; Gilbert, D.J.; Jenkins, N.A.; Copeland, N.G.; Welgus,
J. Biol. Chem. 267, 4664-4671, 1992
A:Title: Molecular cloning, chromosomal localization, and bacterial expression of a m
A:Reference number: A42401; MUID:92165826
A:Accession: A42401
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-462 <SHA>
A:Cross-references: GB:M82831; NID:gl99127; PIDN:AAA39526.1; PID:gl99128
C:Superfamily: interstitial collagenase; hemopexin repeat homology; matrix metallopro
C:Keywords: hydrolase; metalloproteinase; zinc; zymogen
F:53-256/Domain: matrix metalloproteinase homology <WMP>
F:269-462/Domain: hemopexin repeat homology <PXN>
F:85,211,215,221/Binding site: zinc, catalytic (Cys, His, His) (inhibited) #stat
F:211,215,221/Binding site: zinc, catalytic (His) (active) #status predicted
F:212/Active site: Glu #status predicted

Query Match 75.0%; Score 33; DB 2; Length 462;
Best Local Similarity 83.3%; Pred. No. 63;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMVPTY 6
: | | | | :
Db 228 IMVPTY 233

RESULT 11
T19496
hypothetical protein C2A7.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 26-May-2000
C:Accession: T19496
R:Harris, B.

submitted to the EMBL Data Library, October 1996

A:Reference number: Z19132
 A:Accession: T19496
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-511 <MIL>
 A:Cross-references: EMBL:Z81041; PIDN:CAR02786.1; GSPDB:GN00023; CESP:C27A7.2
 A:Experimental source: clone C27A7
 C:Genetics:
 A:Gene: CESP:C27A7.2
 A:Map position: 5
 A:Introns: 44/3; 87/2; 130/2; 164/3; 203/2; 273/3; 325/1; 369/2
 C:Superfamily: Caenorhabditis elegans hypothetical protein F07G11.3

Query Match 75.0%; Score 33; DB 2; Length 511;
 Best Local Similarity 75.0%; Pred. No. 70;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OK 1 LMPTYIK 8
 291 LMPTYILE 298

RESULT 12
 S69693
 Hypothetical protein YDR401w - yeast (Saccharomyces cerevisiae)
 C:Species: Saccharomyces cerevisiae
 C:Date: 22-Aug-1996 #sequence_revision 06-Sep-1996 #text_change 29-Oct-1999
 C:Accession: S69693
 R:Dieterich, F.S.
 submitted to the EMBL Data Library, July 1995
 A:Description: The sequence of S. cerevisiae cosmid 9481, 9509, 9926, 9461, and lambda
 A:Reference number: S69665
 A:Accession: S69693
 A:Molecule type: DNA
 A:Residues: 1-187 <DIE>
 A:Cross-references: EMBL:U32274; NID:g927313; PID:g2194163; GSPDB:GN00004; MIPS:YDR401w
 C:Genetics:
 A:Gene: MIPS:YDR401w
 A:Map position: 4R
 C:Superfamily: Saccharomyces hypothetical protein YDR401w

Query Match 72.7%; Score 32; DB 2; Length 187;
 Best Local Similarity 83.3%; Pred. No. 40;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OK 1 LMPTY 6
 112 LMPTY 117

RESULT 13
 KCHUM
 matrilysin (EC 3.4.24.23) precursor - human
 N:Alternate names: matrin; matrix metalloproteinase 7 (MMP7); probable metalloproteinase
 N:Contains: promatrilysin
 C:Species: Homo sapiens (man)
 C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 18-Jun-1999
 A:Accession: B28816; A60539; S24324
 R:Miller, D.; Quantin, B.; Gesnel, M.C.; Millon-Collard, R.; Abecassis, J.; Breathnach, Biochem. J. 253, 187-192, 1988
 A:Title: The collagenase gene family in humans consists of at least four members.
 A:Reference number: A90339; MUID:88339885
 A:Accession: B28816
 A:Molecule type: mRNA
 A:Residues: 1-267 <MUL>
 A:Cross-references: EMBL:X07819; NID:g35798; PIDN:CAA30678.1; PID:g35799
 R:Miyazaki, K.; Hattori, Y.; Umenishi, F.; Yasumitsu, H.; Umeda, M. Cancer Res. 50, 7758-7764, 1990
 A:Title: Purification and characterization of extracellular matrix-degrading metalloprote
 A:Reference number: A60539; MUID:91070531
 A:Accession: A60539

A:Molecule type: protein
 A:Residues: 18-35, X, 37-42 <MI>
 R:Marti, H.P.; McNeil, L.; Thomas, G.; Davies, M.; Lovett, D.H. Biochem. J. 285, 899-905, 1992
 A:Title: Molecular characterization of a low-molecular-mass matrix metalloproteinase
 A:Reference number: S24324; MUID:92359961
 A:Accession: S24324
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-267 <MAR>
 A:Cross-references: EMBL:Z11887; NID:g35802; PIDN:CAA77942.1; PID:g35803
 C:Comment: This enzyme is similar in its activity to stromelysin and degrades various
 s of types II, IV, IX, X, and XI
 C:Comment: Matrilysin hydrolyzes peptide bonds in plasminogen to yield a fragment wit
 C:Genetics:
 A:Gene: GDB:MMP7; MFSLI
 A:Cross-references: GDB:125751; OMIM:178990
 A:Map position: 11q21-11q22
 C:Superfamily: matrilysin; matrix metalloproteinase homology
 C:Keywords: calcium; extracellular matrix; fibroblast; hydrolase; metalloproteinase;
 F:1-17/Domain: signal sequence #status predicted <SIG>
 F:18-267/Product: promatrilysin #status predicted <PRO>
 F:18-94/Domain: activation peptide #status predicted <ACT>
 F:55-259/Domain: matrix metalloproteinase homology <MMP>
 F:85-92/Region: autoinhibitory
 F:95-267/Product: matrilysin #status predicted <MAT>
 F:87,214,218,224/Binding site: zinc, catalytic (Cys, His, His, His) (inhibited) #stat
 F:214,218,224/Binding site: zinc, catalytic (His) (active) #status predicted
 F:215/Active site: Glu #status predicted

Query Match 72.7%; Score 32; DB 1; Length 267;
 Best Local Similarity 83.3%; Pred. No. 57;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OK 1 LMPTY 6
 Db 231 VMPTY 236

RESULT 14
 A57490
 matrilysin (EC 3.4.24.23) precursor - rat
 N:Alternate names: matrix metalloproteinase 7 (MMP7)
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 08-Dec-1995 #sequence_revision 08-Dec-1995 #text_change 22-Jun-1999
 C:Accession: A57490
 R:Abramson, S.R.; Conner, G.E.; Nagase, H.; Neuhaus, I.; Woessner Jr., J.F. J. Biol. Chem. 270, 16016-16022, 1995
 A:Title: Characterization of rat uterine matrilysin and its cDNA. Relationship to hum
 A:Reference number: A57490; MUID:95332299
 A:Accession: A57490
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-267 <BB>
 A:Cross-references: GB:L24374; NID:g402492; PIDN:AAA99432.1; PID:g402493
 C:Superfamily: matrilysin; matrix metalloproteinase homology
 C:Keywords: hydrolase; metalloproteinase; zinc; zymogen
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-267/Product: matrilysin #status predicted <MAT>
 F:58-262/Domain: matrix metalloproteinase homology <MMP>
 F:90,217,221,227/Binding site: zinc, catalytic (Cys, His, His, His) (inhibited) #stat
 F:217,221,227/Binding site: zinc, catalytic (His) (active) #status predicted
 F:218/Active site: Glu #status predicted

Query Match 72.7%; Score 32; DB 2; Length 267;
 Best Local Similarity 83.3%; Pred. No. 57;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OK 1 LMPTY 6
 Db 234 VMPTY 239

RESULT 15

JU0458
peroxidase (EC 1.11.1.7) E - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 04-Mar-2000
C:Accession: JU0458
R:Intapruk, C.; Higashimura, N.; Yamamoto, K.; Okada, N.; Shinmyo, A.; Takano, M.
Gene 98, 237-241, 1991
A:Title: Nucleotide sequences of two genomic DNAs encoding peroxidase of Arabidopsis thaliana
A:Reference number: JU0457; MUID:91200671
A:Accession: JU0458
A:Molecule type: DNA
A:Residues: 1-349 <INT>
A:Cross-references: GB:M58381; NID:g166806; PIDN:AAA32842.1; PID:g166807
C:Genetics:
A:Gene: prxEa
A:Introns: 76/3; 140/3; 197/1
C:Superfamily: peroxidase
Keywords: chromoprotein; heme; iron; metalloprotein; oxidoreductase
F:40-120/Disulfide bonds: #status predicted
F:67/Active site: Arg #status predicted
F:71,199/Binding site: heme iron (His) (axial ligands) #status predicted
F:73-78/Disulfide bonds: #status predicted
F:126-329/Disulfide bonds: #status predicted
F:206-238/Disulfide bonds: #status predicted

Query Match 72.7% Score 32; DB 2; Length 349;
Best Local Similarity 83.3% Pred. No. 75;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYL 7
:|||||
Db 226 LYPTYL 231

Search completed: September 4, 2001, 15:52:08
Job time: 231 sec

THIS PAGE BLANK (uspto)

GenCoré version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2001, 15:49:03 ; Search time 27.19 Seconds
(without alignments)
10.079 Million cell updates/sec

Title: US-09-630-345-1
Perfect score: 44
Sequence: 1 LMVPTYLK 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41	93.2	384	MM01_RANCA	Q11133 rana catesb
2	36	81.8	340	CELF_VZVD	P09261 varicella-z
3	36	81.8	436	GABT_BACSU	P94427 bacillus su
4	35	79.5	467	MM18_XENLA	Q13065 xenopus lae
5	35	79.5	482	MM20_MOUSE	P57748 mus musculus
6	35	79.5	483	MM20_HUMAN	O60882 homo sapien
7	35	79.5	483	MM20_PIG	P79287 sus scrofa
8	33	75.0	462	MM12_MOUSE	P34960 mus musculus
9	33	75.0	465	MM12_RAT	O63341 rattus norv
10	33	75.0	511	YELU_CAEEL	P90756 caenorhabdi
11	32	72.7	262	MM07_FELCA	P55032 felis silve
12	32	72.7	264	MM07_MOUSE	Q10738 mus musculus
13	32	72.7	267	MM07_HUMAN	P09237 homo sapien
14	32	72.7	267	MM07_RAT	P50280 rattus norv
15	32	72.7	349	PERE_ARATH	P24102 arabidopsis
16	32	72.7	468	MM01_RABIT	P13943 oryctolagus
17	32	72.7	469	MM01_PIG	P21692 sus scrofa
18	32	72.7	594	NUSM_HIPAM	Q92zy1 hippopotamu
19	32	72.7	1263	RPOB_THEMA	P29398 thermotoga
20	31	70.5	301	SC14_CANAL	P46250 candida alb
21	31	70.5	402	OXAL1_YEAST	P39952 saccharomyc
22	31	70.5	423	BIOA_SERMA	P36568 serratia ma
23	31	70.5	469	MM01_BOVIN	P28053 bos taurus
24	31	70.5	469	MM01_HUMAN	P03956 homo sapien
25	31	70.5	481	MM20_BOVIN	O18767 bos taurus
26	31	70.5	495	VLL1_BPV1	P03103 bovine papi
27	31	70.5	497	VLL1_BPV2	P06458 bovine papi
28	31	70.5	501	VLL1_PAPVE	P11326 bovine el
29	31	70.5	503	YPTJ_CAEEL	Q23469 caenorhabdi
30	31	70.5	505	TUB_MOUSE	P50586 mus musculus
31	31	70.5	505	TUB_RAT	O88808 rattus norv
32	31	70.5	506	TUB_HUMAN	P50607 homo sapien
33	31	70.5	513	VLL1_PAPVD	P03104 deer papill

34	31	70.5	662	1	YME1_SCHMA	P46508 schistosoma
35	31	70.5	675	1	KSC5_ECOLI	P42217 escherichia
36	31	70.5	1068	1	YCF0_MARPO	P12221 marchantia
37	31	70.5	1440	1	SYEP_HUMAN	P07814 homo sapien
38	31	70.5	1714	1	SYEP_DROME	P28668 drosophila
39	31	70.5	2185	1	POLG_SVDVH	P16604 s genome po
40	31	70.5	2185	1	POLG_SVDVU	P13900 s genome po
41	30	68.2	91	1	X13F_BPT4	P39499 bacterioph
42	30	68.2	124	1	FLGB_AQUAE	O67244 aquifex aeo
43	30	68.2	192	1	NUMG_PROWI	Q37622 escherichia
44	30	68.2	239	1	YKGE_ECOLI	P77252 escherichia
45	30	68.2	305	1	NUMM_APILI	P34847 apis mellif

ALIGNMENTS

RESULT 1	MM01_RANCA	STANDARD;	PRT;	384 AA.
ID	MM01_RANCA	STANDARD;	PRT;	384 AA.
AC	Q11133;			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	01-OCT-2000 (Rel. 40, Last annotation update)			
DE	INTERSTITIAL COLLAGENASE PRECURSOR (EC 3.4.24.7) (MATRIX METALLOPROTEINASE-1) (MMP-1) (YCL).			
OS	Rana catesbeiana (Bull frog).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Amphibia; Batrachia; Anura; Neobatrachia; Ranidae; Rana.			
OX	NCBI_TaxID=8400;			
[1]				
RP	SEQUENCE FROM N.A.			
TX	TISSUE=Skin;			
RC	MEDLINE=95071832; PubMed=7981043;			
RA	Oofusa K., Yomori S., Yoshizato K.;			
RT	"Regionally and hormonally regulated expression of genes of collagen and collagenase in the anuran larval skin.";			
RL	Int. J. Dev. Biol. 38:345-350(1994).			
CC	-!- FUNCTION: CLEAVES COLLAGENS OF TYPES I, II, AND III AT ONE SITE IN THE HELICAL DOMAIN. ALSO CLEAVES COLLAGENS OF TYPES VII AND X.			
CC	-!- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY.			
CC	-!- ENZYME REGULATION: CAN BE ACTIVATED WITHOUT REMOVAL OF THE ACTIVATION PEPTIDE (BY SIMILARITY).			
CC	-!- SIMILARITY: CONTAINS 1 HEMOPEXIN-LIKE DOMAIN.			
CC	-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC METALLOPROTEASE) ALSO KNOWN AS MATRIXIN SUBFAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
EMBL	S75623; AAB32661.1; -			
HSSP	P21692; IFBL.			
DR	MEROPS; M10.001; -			
DR	InterPro; IPR000130; -			
DR	InterPro; IPR000585; -			
DR	InterPro; IPR001818; -			
DR	Pfam; PF00413; Peptidase_M10; 1.			
DR	Pfam; PF00045; Hemoexin; 2.			
DR	PROSITE; PS00024; HEMOPEXIN; FALSE_NEG.			
DR	PROSITE; PS00142; ZINC_PROTEASE; 1.			
DR	PROSITE; PS00546; CYSTEINE_SWITCH; FALSE_NEG.			
DR	Hydrolase; Metalloprotease; Zinc; Zymogen; Calcium;			
KW	Collagen degradation; Extracellular matrix; Signal.			
FT	SIGNAL 1 25 POTENTIAL.			
FT	PROPEP 26 88 ACTIVATION PEPTIDE (POTENTIAL).			
FT	CHAIN 89 384 INTERSTITIAL COLLAGENASE.			
FT	DOMAIN 239 384 HEMOPEXIN-LIKE.			
FT	SITE 81 81 CYSTEINE SWITCH (POTENTIAL).			

```

FT METAL 189 189 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 190 190 BY SIMILARITY.
FT METAL 193 193 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 199 199 ZINC (CATALYTIC) (BY SIMILARITY).
FT DISULFID 242 381 PROBABLE.
SQ SEQUENCE 384 AA; 43582 MW; A5B5E2FB332239DF CRC64;

Query Match 93.2%; Score 41; DB 1; Length 384;
Best Local Similarity 87.5%; Pred. No. 0.82;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LMPTYTLK 8
Db 206 LMPTYTLR 213

RESULT 2
GABT_VZVD STANDARD; PRT; 340 AA.
P03261:
01-MAR-1989 (Rel. 10, Created)
01-MAR-1989 (Rel. 10, Last sequence update)
15-DEC-1998 (Rel. 37, Last annotation update)
DE CELL FUSION PROTEIN PRECURSOR.
GN 5.
OS Varicella-zoster virus (strain Dumas) (VZV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10338;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86306657; PubMed=3018124;
RA Davison A.J., Scott J.E.;
RT "The complete DNA sequence of varicella-zoster virus.";
RL J. Gen. Virol. 67:1759-1816(1986).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
DR EMBL; X04370; CAA27888.1; -.
DR F01621; FUSION_GLY_K; 1.
DR InterPro: IPR002567; -.
KW Fusion protein; Transmembrane; Signal.
FT SIGNAL 1 ?
FT CHAIN ? 340 CELL FUSION PROTEIN.
FT CARBOHYD 55 55 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 65 65 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 340 AA; 38576 MW; 0387FE08C39C946 CRC64;

Query Match 81.8%; Score 36; DB 1; Length 340;
Best Local Similarity 62.5%; Pred. No. 6.6;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LMPTYTLK 8
Db 243 LLYPTYIR 250

RESULT 3
GABT_BACSU STANDARD; PRT; 436 AA.
AC P94427;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)

```

```

DE PROBABLE 4-AMINOBUTYRATE AMINOTRANSFERASE (EC 2.6.1.19) (GAMMA-AMINO-
DE N-BUTYRATE TRANSAMINASE) (GABA TRANSAMINASE) (GLUTAMATE:SUCCINIC
DE SEMIALDEHYDE TRANSAMINASE) (GABA AMINOTRANSFERASE) (GABA-AT).
GN GABT.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=168;
RX MEDLINE=97124189; PubMed=8969502;
RA Yamane K., Kumano M., Kurita K.;
RT "The 25 degrees-36 degrees region of the Bacillus subtilis
RT chromosome: determination of the sequence of a 146 kb segment and
RT identification of 113 genes."
RL Microbiology 142:3047-3056(1996).
CC -!- CATALYTIC ACTIVITY: 4-AMINOBUTANOATE + 2-OXOGLUTARATE -> SUCCINATE
CC SEMIALDEHYDE + L-GLUTAMATE.
CC -!- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -!- PATHWAY: 4-AMINOBUTYRATE (GABA) DEGRADATION PATHWAY.
CC -!- SIMILARITY: BELONGS TO CLASS-III OF PYRIDOXAL-PHOSPHATE-DEPENDENT
CC AMINOTRANSFERASES.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
DR EMBL; D50453; BAA09021.1; -.
DR EMBL; Z99106; CAB12198.1; -.
DR Subtilist; BG12043; gabt.
DR InterPro: IPR000954; -.
DR Pfam: PF00202; aminotran_3; 1.
DR PROSITE: PS00600; AA-TRANSFER CLASS_3; 1.
KW Transferase; Aminotransferase; Pyridoxal phosphate.
FT BINDING 281 281 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ SEQUENCE 436 AA; 47249 MW; D0961F6DA189A8F3 CRC64;

Query Match 81.8%; Score 36; DB 1; Length 436;
Best Local Similarity 62.5%; Pred. No. 8.5;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LMPTYTLK 8
Db 87 NMYPTYIE 94

RESULT 4
MM18_XENLA STANDARD; PRT; 467 AA.
AC O13065;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE MATRIX METALLOPROTEINASE-18 PRECURSOR (EC 3.4.24.-) (MMP-18)
DE (COLLAGENASE-4) (COLLAGENASE 4) (XCOL4).
GN MMP18 OR COL4.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Intestine;
RX MEDLINE=97053976; PubMed=8898355;
RA Stelow M.A., Bauzon D.D., Li J., Sedgwick T., Liang V.C.-T.,
RA Sang Q.A., Shi Y.-B.;

```

RT Identification and characterization of a novel collagenase in Xenopus
 RL laevis: possible roles during frog development.";
 CC Mol. Biol. Cell 7:1471-1483(1996).
 CC -1- FUNCTION: CLEAVES COLLAGEN TYPE I. MAY PLAY A ROLE IN LARVAL
 CC TISSUE DEGENERATION AND ADULT ORGANOGENESIS DURING AMPHIBIAN
 CC METAMORPHOSIS. MAY BE INVOLVED IN TAIL RESORPTION.
 CC -1- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY (BY SIMILARITY).
 CC -1- ENZYME REGULATION: UPREGULATED IN THE TAIL BY THYROID HORMONE.
 CC -1- SUBCELLULAR LOCATION: SECRETED. ASSOCIATED WITH THE EXTRACELLULAR
 CC MATRIX (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: EXPRESSED ONLY TRANSIENTLY IN WHOLE ANIMAL, AT
 CC TIME WHEN TADPOLE FEEDING BEGINS.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED AT HIGH LEVEL AS THE TADPOLE TAIL
 CC RESORBS AND DURING HINDLIMB MORPHOGENESIS AND INTESTINAL
 CC REMODELING.
 CC -1- SIMILARITY: CONTAINS 1 HEMOPEXIN-LIKE DOMAIN.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC
 CC METALLOPROTEASE) ALSO KNOWN AS MATRIXIN SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L76275; AAB53148.1; -;
 CC HSSP; P03956; 1CGL.
 CC MEROPS; M10.018; -;
 CC InterPro; IPR000130; -;
 CC InterPro; IPR000585; -;
 CC InterPro; IPR001818; -;
 CC Pfam; PF00045; hemopexin; 4.
 CC Pfam; PF00413; Peptidase.M10; 1.
 CC PRINTS; PR00138; MATRIXIN.
 CC PROSITE; PS00142; ZINC_PROTEASE; 1.
 CC PROSITE; PS00546; CYSTEINE_SWITCH; FALSE_NEG.
 CC PROSITE; PS00024; HEMOPEXIN; FALSE_NEG.
 CC KW Hydrolyase; Metalloprotease; Zinc; Calcium; Zymogen; Signal;
 CC Collagen degradation; Extracellular matrix.
 CC SIGNAL 1 17 POTENTIAL.
 CC FT PROPEP 18 99 BY SIMILARITY.
 CC FT CHAIN 100 467 MATRIX METALLOPROTEINASE-18.
 CC FT DOMAIN 277 467 HEMOPEXIN-LIKE.
 CC FT SITE 92 92 CYSTEINE SWITCH (POTENTIAL).
 CC FT METAL 218 218 ZINC (CATALYTIC) (BY SIMILARITY).
 CC FT ACT_SITE 219 219 BY SIMILARITY.
 CC FT METAL 222 222 ZINC (CATALYTIC) (BY SIMILARITY).
 CC FT METAL 228 228 ZINC (CATALYTIC) (BY SIMILARITY).
 CC FT DISULFID 280 467 POTENTIAL.
 CC SQ SEQUENCE 467 AA; 52812 MW; 4623F6CEFF3454051 CRC64;
 Query Match 79.5%; Score 35; DB 1; Length 467;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LMVPTY 6
 | | | | |
 DB 235 LMVPTY 240
 RESULT 5
 MM20_MOUSE
 ID MM20_MOUSE STANDARD; PRT; 482 AA.
 AC P57748;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE MATRIX METALLOPROTEINASE-20 PRECURSOR (EC 3.4.24.-) (MMP-20) (ENAMEL
 DE METALLOPROTEINASE) (ENAMELYSIN).
 GN MMP20.

OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1].
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/SVJ;
 RX MEDLINE=20079167; PubMed=10610728;
 RA Caterina J., Shi J., Krakora S., Bartlett J.D., Engler J.A.,
 RA Kozak C.A., Birkedal-Hansen H.;
 RT "Isolation, characterization, and chromosomal location of the mouse
 RT enamelysin gene.";
 RL Genomics 62:308-311(1999).
 CC -1- FUNCTION: DEGRADES AMELOGENIN, THE MAJOR PROTEIN COMPONENT OF THE
 CC ENAMEL MATRIX AND TWO OF THE MACROMOLECULES CHARACTERISING THE
 CC CARTILAGE EXTRACELLULAR MATRIX: AGGRECAN AND THE CARTILAGE
 CC OLIGOMERIC MATRIX PROTEIN (COMP). MAY PLAY A CENTRAL ROLE IN TOOTH
 CC ENAMEL FORMATION (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: CLEAVES AGGRECAN AT THE 360-ASN-1-PHE-361 SITE
 CC (BY SIMILARITY).
 CC -1- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: SECRETED. ASSOCIATED WITH THE EXTRACELLULAR
 CC MATRIX (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: EXPRESSED SPECIFICALLY IN THE ENAMEL ORGAN.
 CC -1- PTM: AUTOACTIVATES AT LEAST AT THE 106-ASN-1-TYR-107 SITE (BY
 CC SIMILARITY).
 CC -1- SIMILARITY: CONTAINS 1 HEMOPEXIN-LIKE DOMAIN.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC
 CC METALLOPROTEASE) ALSO KNOWN AS MATRIXIN SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF156956; AAF28472.1; -;
 CC EMBL; AF156947; AAF28472.1; JOINED.
 CC EMBL; AF156948; AAF28472.1; JOINED.
 CC EMBL; AF156949; AAF28472.1; JOINED.
 CC EMBL; AF156950; AAF28472.1; JOINED.
 CC EMBL; AF156951; AAF28472.1; JOINED.
 CC EMBL; AF156952; AAF28472.1; JOINED.
 CC EMBL; AF156953; AAF28472.1; JOINED.
 CC EMBL; AF156954; AAF28472.1; JOINED.
 CC EMBL; AF156955; AAF28472.1; JOINED.
 CC EMBL; AF155933; AAF28470.1; -;
 CC MEROPS; M10.019; -;
 CC MGD; MGI:1353466; Mmp20.
 CC PROSITE; PS00546; CYSTEINE_SWITCH; 1.
 CC PROSITE; PS00142; ZINC_PROTEASE; 1.
 CC PROSITE; PS00024; HEMOPEXIN; FALSE_NEG.
 CC KW Hydrolyase; Metalloprotease; Zinc; Calcium; Zymogen; Signal;
 CC Extracellular matrix.
 CC SIGNAL 1 21 POTENTIAL.
 CC FT PROPEP 22 106 BY SIMILARITY.
 CC FT CHAIN 107 482 MATRIX METALLOPROTEINASE-20.
 CC FT DOMAIN 292 482 HEMOPEXIN-LIKE.
 CC FT SITE 99 99 CYSTEINE SWITCH (POTENTIAL).
 CC FT METAL 225 225 ZINC (CATALYTIC) (BY SIMILARITY).
 CC FT ACT_SITE 226 226 BY SIMILARITY.
 CC FT METAL 229 229 ZINC (CATALYTIC) (BY SIMILARITY).
 CC FT METAL 235 235 ZINC (CATALYTIC) (BY SIMILARITY).
 CC FT DISULFID 295 482 BY SIMILARITY.
 CC SQ SEQUENCE 482 AA; 54373 MW; 366149FAF2BDC8BB CRC64;
 Query Match 79.5%; Score 35; DB 1; Length 482;
 Best Local Similarity 100.0%; Pred. No. 15;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 LMPTY 6
DB 242 LMPTY 247

RESULT 6
MM20_HUMAN STANDARD; PRT; 483 AA.
AC OG0882;
DT 01-OCT-2000 (rel. 40, Created)
DT 01-OCT-2000 (rel. 40, Last sequence update)
DT 01-OCT-2000 (rel. 40, Last annotation update)
DE MATRIX METALLOPROTEINASE-20 PRECURSOR (EC 3.4.24.-) (MMP-20) (ENAMEL
DE METALLOPROTEINASE) (ENAMELYSIN).
GN MMP20.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A., AND FUNCTION.
RC TISSUE=Odontoblast;
RX MEDLINE=98060762; PubMed=9398237;
RA Llano E., Pendas A.M., Knaeuper V., Sorsa T., Salo T., Salido E.,
RA Murphy G., Simmer J.P., Bartlett J.D., Lopez-Otin C.;
RT "Identification and structural and functional characterization of
RT human enamelysin (MMP-20).";
RL Biochemistry 36:15101-15108(1997).
RN [2]
FUNCTION.
RX MEDLINE=20382723; PubMed=10922468;
RA Stracke J.O., Fosang A.J., Last K., Mercuri F.A., Pendas A.M.,
RA Llano E., Perris R., Di Cesare P.E., Murphy G., Knaeuper V.;
RA "Matrix metalloproteinases 19 and 20 cleave aggrecan and cartilage
RT oligomeric matrix protein (COMP).";
RL FEBS Lett. 478:52-56(2000).
RN [2]
FUNCTION.
CC -!- FUNCTION: DEGRADATES AMELOGENIN, THE MAJOR PROTEIN COMPONENT OF THE
CC ENAMEL MATRIX AND TWO OF THE MACROMOLECULES CHARACTERISING THE
CC CARTILAGE EXTRACELLULAR MATRIX: AGGREGAN AND THE CARTILAGE
CC OLIGOMERIC MATRIX PROTEIN (COMP). MAY PLAY A CENTRAL ROLE IN TOOTH
CC ENAMEL FORMATION.
CC -!- CATALYTIC ACTIVITY: CLEAVES AGGREGAN AT THE 360-ASN-|-PHE-361
CC SITE.
CC -!- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY.
CC -!- SUBCELLULAR LOCATION: SECRETED. ASSOCIATED WITH THE EXTRACELLULAR
CC MATRIX (BY SIMILARITY).
CC -!- TISSUE SPECIFICITY: EXPRESSED SPECIFICALLY IN THE ENAMEL ORGAN.
CC -!- PTM: AUTOACTIVATES AT LEAST AT THE 107-ASN-|-TYR-108 SITE (BY
CC -!- SIMILARITY).
CC -!- SIMILARITY: CONTAINS 1 HEMOPEXIN-LIKE DOMAIN.
CC -!- METALLOPROTEINASE) ALSO KNOWN AS MATRIXIN SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Y12779; CAA73317.1;
CC DR MEROPS: M10.019;
CC DR HSSP: P08254; LUSN
CC DR InterPro: IPR000130;
CC DR InterPro: IPR000585;
CC DR Pfam: PF000413; hemopexin; 4.
CC DR PRINTS: PR00138; Peptidase_M10; 1.
CC DR PROSITE: PS00546; CYSTEINE_SWITCH; 1.
CC DR PROSITE: PS00024; HEMOPEXIN; FALSE_NEG

PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Hydrolase; Metalloprotease; Zinc; Calcium; Zymogen; Signal;
KW Extracellular matrix.
FT SIGNAL 1 22 POTENTIAL.
FT PROPEP 23 107 BY SIMILARITY.
FT CHAIN 108 483 MATRIX METALLOPROTEINASE-20.
FT DOMAIN 293 483 HEMOPEXIN-LIKE.
FT SITE 100 100 CYSTEINE SWITCH (POTENTIAL).
FT METAL 226 226 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 227 227 BY SIMILARITY.
FT METAL 230 230 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 236 236 ZINC (CATALYTIC) (BY SIMILARITY).
FT DISULFID 296 483 POTENTIAL.
FT SEQUENCE 483 AA; 54413 MW; 1DABBA04967AB395 CRC64;

Query Match 79.5%; Score 35; DB 1; Length 483;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMPTY 6
DB 243 LMPTY 248

RESULT 7
MM20_PIG STANDARD; PRT; 483 AA.
AC P78287;
DT 01-OCT-2000 (rel. 40, Created)
DT 01-OCT-2000 (rel. 40, Last sequence update)
DT 01-OCT-2000 (rel. 40, Last annotation update)
DE MATRIX METALLOPROTEINASE-20 PRECURSOR (EC 3.4.24.-) (MMP-20) (ENAMEL
DE METALLOPROTEINASE) (ENAMELYSIN).
GN MMP20.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
NCBI_TaxID=9823;
[1]
SEQUENCE FROM N.A.
RC TISSUE=Enamel organ;
RX MEDLINE=97149288; PubMed=8996096;
RA Bartlett J.D., Simmer J.P., Xue J., Margolis H.C., Moreno E.C.;
RT "Molecular cloning and mRNA tissue distribution of a novel matrix
RT metalloproteinase isolated from porcine enamel organ.";
RL Gene 183:123-128(1996).
RN [1]
FUNCTION: DEGRADATES AMELOGENIN, THE MAJOR PROTEIN COMPONENT OF THE
CC ENAMEL MATRIX AND TWO OF THE MACROMOLECULES CHARACTERISING THE
CC CARTILAGE EXTRACELLULAR MATRIX: AGGREGAN AND THE CARTILAGE
CC OLIGOMERIC MATRIX PROTEIN (COMP). MAY PLAY A CENTRAL ROLE IN TOOTH
CC ENAMEL FORMATION (BY SIMILARITY).
CC -!- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY.
CC -!- SUBCELLULAR LOCATION: SECRETED. ASSOCIATED WITH THE EXTRACELLULAR
CC MATRIX (BY SIMILARITY).
CC -!- TISSUE SPECIFICITY: EXPRESSED SPECIFICALLY IN THE ENAMEL ORGAN.
CC -!- PTM: AUTOACTIVATES AT LEAST AT THE 107-ASN-|-TYR-108 SITE (BY
CC -!- SIMILARITY).
CC -!- SIMILARITY: CONTAINS 1 HEMOPEXIN-LIKE DOMAIN.
CC -!- METALLOPROTEINASE) ALSO KNOWN AS MATRIXIN SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U54825; AAB41396.1;
CC DR HSSP: P03956; ACGL.
CC DR MEROPS: M10.019;

```


DR InterPro: IPR000130; -
 DR InterPro: IPR000585; -
 DR InterPro: IPR001818; -
 DR Pfam: PF00045; hemopexin; 4.
 DR Pfam: PF00413; peptidase_M10; 1.
 DR PRINTS: PR00138; MATRINXIN.
 DR PROSITE: PS00024; HEMOPEXIN; FALSE_NEG.
 DR PROSITE: PS00142; ZINC_PROTEASE; 1.
 DR PROSITE: PS00546; CYSTEINE_SWITCH; 1.
 KW Hydroxylase; Metalloprotease; Zinc; Calcium; Zymogen; Signal;
 KW Extracellular matrix.
 FT SIGNAL 1 22 POTENTIAL.
 FT PROPEP 23 107 ACTIVATION PEPTIDE (POTENTIAL).
 FT CHAIN 108 483 MATRIX METALLOPROTEINASE-20.
 FT DOMAIN 293 483 HEMOPEXIN-LIKE.
 FT SITE 100 100 CYSTEINE SWITCH (POTENTIAL).
 FT METAL 226 226 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACT_SITE 227 227 BY SIMILARITY.
 FT METAL 230 230 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 236 236 ZINC (CATALYTIC) (BY SIMILARITY).
 FT DISULFID 296 483 BY SIMILARITY.
 SEQUENCE 483 AA; 54084 MW; 1295DAAAEFIA20B4 CRC64;

Query Match 79.5%; Score 35; DB 1; Length 483;
 Best Local Similarity 100.0%; Pred. No. 15; Mismatches 0; Indels 0; Gaps 0;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMPTY 6
 :|||||
 DB 243 LMPTY 248

RESULT 8
 MM12_MOUSE STANDARD; PRT; 462 AA.
 AC P34960;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE MACROPHAGE METALLOELASTASE PRECURSOR (EC 3.4.24.65) (MME) (MATRIX METALLOPROTEINASE-12) (MMP-12).
 GN MMP12 OR MME OR MME.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 99-125.
 TT TISSUE=Macrophage;
 MEDLINE=92165826; PubMed=1537850;
 RA Shapiro S.D., Griffin G.L., Gilbert D.J., Jenkins N.A., Copeland N.G., Welgus H.G., Senior R.M., Ley T.J.;
 RT "Molecular cloning, chromosomal localization, and bacterial expression of a murine macrophage metalloelastase.";
 RL J. Biol. Chem. 267:4664-4671(1992).
 CC -1- FUNCTION: MAY BE INVOLVED IN TISSUE INJURY AND REMODELING. HAS SIGNIFICANT ELASTOLYTIC ACTIVITY.
 CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF SOLUBLE AND INSOLUBLE ELASTIN. SPECIFIC CLEAVAGES ARE ALSO PRODUCED AT 14-ALA-|-LEU-15 AND 16-TYR-|-LEU-17 IN THE B CHAIN OF INSULIN.
 CC -1- COPACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY.
 CC -1- SIMILARITY: CONTAINS 1 HEMOPEXIN-LIKE DOMAIN.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC METALLOPROTEASE) ALSO KNOWN AS MATRINXIN SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: M82831; AAA39526.1; -
 DR PIR: A42401; A42401.
 DR HSP: P03956; ICGL.
 DR MEROPS: M10.009; -
 DR MGD: MGI:97005; Mmp12.
 DR InterPro: IPR000130; -
 DR InterPro: IPR000585; -
 DR InterPro: IPR001818; -
 DR Pfam: PF00413; peptidase_M10; 1.
 DR Pfam: PF00045; hemopexin; 4.
 DR PRINTS: PR00138; MATRINXIN.
 DR PROSITE: PS00024; HEMOPEXIN; 1.
 DR PROSITE: PS00142; ZINC_PROTEASE; 1.
 DR PROSITE: PS00546; CYSTEINE_SWITCH; 1.
 KW Hydroxylase; Metalloprotease; Glycoprotein; Zinc; Zymogen; Calcium;
 KW Extracellular matrix; Signal.
 FT SIGNAL 1 17 PROBABLE.
 FT PROPEP 18 98 ACTIVATION PEPTIDE.
 FT CHAIN 99 462 MACROPHAGE METALLOELASTASE.
 FT DOMAIN 272 462 HEMOPEXIN-LIKE.
 FT SITE 85 85 CYSTEINE SWITCH (BY SIMILARITY).
 FT METAL 211 211 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACT_SITE 212 212 BY SIMILARITY.
 FT METAL 215 215 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 221 221 ZINC (CATALYTIC) (BY SIMILARITY).
 FT CARBOHYD 21 21 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 74 74 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 310 310 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT DISULFID 275 462 BY SIMILARITY.
 SEQUENCE 462 AA; 53841 MW; BB9625906F1DBEDF CRC64;

Query Match 75.0%; Score 33; DB 1; Length 462;
 Best Local Similarity 83.3%; Pred. No. 34;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMPTY 6
 :|||||
 DB 228 LMPTY 233

RESULT 9
 MM12_RAT STANDARD; PRT; 465 AA.
 ID AC Q63341;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE MACROPHAGE METALLOELASTASE PRECURSOR (EC 3.4.24.65) (MME) (MATRIX METALLOPROTEINASE-12) (MMP-12).
 GN MMP12 OR MMEL.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY;
 RA Cossins J., Clements J., Catlin G.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: MAY BE INVOLVED IN TISSUE INJURY AND REMODELING. HAS SIGNIFICANT ELASTOLYTIC ACTIVITY (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF SOLUBLE AND INSOLUBLE ELASTIN. SPECIFIC CLEAVAGES ARE ALSO PRODUCED AT 14-ALA-|-LEU-15 AND 16-TYR-|-LEU-17 IN THE B CHAIN OF INSULIN.
 CC -1- COPACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY (BY SIMILARITY).
 CC -1- SIMILARITY: CONTAINS 1 HEMOPEXIN-LIKE DOMAIN.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC METALLOPROTEASE) ALSO KNOWN AS MATRINXIN SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

FT METAL 219 219 ZINC (CATALYTIC) (BY SIMILARITY).
SQ SEQUENCE 262 AA; 29263 MW; E4A1FA23320DC732 CRC64;

Query Match 72.7%; Score 32; DB 1; Length 262;
Best Local Similarity 83.3%; Pred. No. 30;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMYPTY 6
:|||||
Db 226 VMVPTY 231

RESULT 12

MM07_MOUSE STANDARD; PRT; 264 AA.
ID MM07_MOUSE Q10738;
AC Q10738;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE MATRILYSIN PRECURSOR (EC 3.4.24.23) (PUMP-1 PROTEASE) (UTERINE METALLOPROTEINASE) (MATRIX METALLOPROTEINASE-7) (MMP-7) (MATRIN).
NMMP7.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR; TISSUE=Uterus;
RX MEDLINE=96086498; PubMed=7579699;
RA Wilson C.L., Heppner K.J., Rudolph L.A., Matrisian L.M.;
RT "The metalloproteinase matrilysin is preferentially expressed by epithelial cells in a tissue-restricted pattern in the mouse.";
RL Mol. Biol. Cell 6:851-869(1995).
CC -1- FUNCTION: DEGRADES CASEIN, GELATINS OF TYPES I, III, IV, AND V, AND FIBRONECTIN. ACTIVATES PROCOLLAGENASE (BY SIMILARITY).
CC -1- FUNCTION: MAY PLAY A ROLE IN TISSUE REORGANIZATION.
CC -1- CATALYTIC ACTIVITY: CLEAVAGE OF 14-ALA-|-LEU-15 AND 16-TYR-|-LEU-17 IN B CHAIN OF INSULIN. NO ACTION ON COLLAGEN TYPES I, II, IV, V. CLEAVES GELATIN CHAIN ALPHA-2(I) > ALPHA-1(I).
CC -1- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC METALLOPROTEASES) ALSO KNOWN AS MATRIXIN SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L36238; AAA99984.1; JOINED.
CC EMBL; L36243; AAA99984.1; JOINED.
CC EMBL; L36242; AAA99984.1; JOINED.
CC EMBL; L36241; AAA99984.1; JOINED.
CC EMBL; L36240; AAA99984.1; JOINED.
CC EMBL; L36239; AAA99984.1; JOINED.
CC EMBL; L36244; AAA99984.1; JOINED.
CC HSP; P09237; 1MMR.
CC MEROPS; M10.008; -.
CC DR MGD; MGI:103189; Mmp7.
CC InterPro; IPR000130; -.
CC InterPro; IPR001818; -.
CC Pfam; PF00413; Peptidase_M10; 1.
CC PRINTS; PR00138; MATRINX.
CC PROSITE; PS00142; ZINC_PROTEASE; 1.
CC PROSITE; PS00546; CYSTEINE_SWITCH; 1.
CC KW Hydrolase; Metalloprotease; Zinc; Zymogen; Calcium;
KW Collagen degradation; Extracellular matrix; Signal.
FT SIGNAL 1 17 POTENTIAL
FT PROPEP 18 94 ACTIVATION PEPTIDE (BY SIMILARITY).

FT CHAIN 95 264 MATRILYSIN.
FT SITE 87 87 CYSTEINE SWITCH (POTENTIAL).
FT METAL 214 214 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 215 215 BY SIMILARITY.
FT METAL 218 218 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 224 224 ZINC (CATALYTIC) (BY SIMILARITY).
FT CONFLICT 201 201 G -> D (IN AAA99983).
SQ SEQUENCE 264 AA; 29755 MW; EDA31A5EBAC63342 CRC64;

Query Match 72.7%; Score 32; DB 1; Length 264;
Best Local Similarity 83.3%; Pred. No. 31;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMYPTY 6
:|||||
Db 231 VMVPTY 236

RESULT 13

MM07_HUMAN STANDARD; PRT; 267 AA.
ID MM07_HUMAN P09237;
AC P09237;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE MATRILYSIN PRECURSOR (EC 3.4.24.23) (PUMP-1 PROTEASE) (UTERINE METALLOPROTEINASE) (MATRIX METALLOPROTEINASE-7) (MMP-7) (MATRIN).
GN MMP7 OR MPSL1 OR PUMPL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88339885; PubMed=2844164;
RA Muller D., Quantin B., Gesnel M.-C., Millon-Collard R., Abecassis J., Breathnach R.;
RT "The collagenase gene family in humans consists of at least four members.";
RL Biochem. J. 253:187-192(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=92359961; PubMed=1497627;
RA Marti H.P., McNeil L., Thomas G., Davies M., Lovett D.H.;
RT "Molecular characterization of a low-molecular-mass matrix metalloproteinase secreted by glomerular mesangial cells as PUMP-1.";
RL Biochem. J. 285:899-905(1992).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=94124555; PubMed=8294454;
RA Gaire M., Magbanua Z., McDonnell S., McNeil L.B., Lovett D.H., Matrisian L.M.;
RT "Structure and expression of the human gene for the matrix metalloproteinase matrilysin.";
RL J. Biol. Chem. 269:2032-2040(1994).
RN [4]
RP SEQUENCE OF 18-42.
RX MEDLINE=91070531; PubMed=2253219;
RA Miyazaki K., Hattori Y., Umenishi F., Yasumitsu H., Umeda M.;
RT "Purification and characterization of extracellular matrix-degrading metalloproteinase, matrin (pump-1), secreted from human rectal carcinoma cell line.";
RL Cancer Res. 50:7758-7764(1990).
RN [5]
RP FUNCTION.
RX MEDLINE=89375247; PubMed=2550050;
RA Quantin B., Murphy G., Breathnach R.;
RT "Pump-1 cDNA codes for a protein with characteristics similar to those of classical collagenase family members.";
RL Biochemistry 28:5327-5334(1989).

[6]
 RN X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
 RP MEDLINE=95275856; PubMed=7756291;
 RX Browner M.F., Smith W.W., Castelano A.L.;
 RA "Matrilysin-inhibitor complexes: common themes among
 RT metalloproteases";
 RL Biochemistry 34:6602-6610(1995).
 CC -!- FUNCTION: DEGRADATES CASEIN, GELATINS OF TYPES I, III, IV, AND V,
 CC AND FIBRONECTIN. ACTIVATES PROCOLLAGENASE.
 CC -!- CATALYTIC ACTIVITY: CLEAVAGE OF 14-ALA-|-LEU-15 AND 16-TYR-|-
 CC LEU-17 IN B CHAIN OF INSULIN. NO ACTION ON COLLAGEN TYPES I, II,
 CC IV, V. CLEAVES GELATIN CHAIN ALPHA-2(I) > ALPHA-1(I).
 CC -!- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC
 CC METALLOPROTEASE) ALSO KNOWN AS MATRININ SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X07819; CA30678.1; -;
 CC EMBL; Z11887; CA307542.1; -;
 CC EMBL; L22524; AAC37543.1; -;
 CC EMBL; L22519; AAC37543.1; JOINED.
 CC EMBL; L22520; AAC37543.1; JOINED.
 CC EMBL; L22521; AAC37543.1; JOINED.
 CC EMBL; L22522; AAC37543.1; JOINED.
 CC EMBL; L22523; AAC37543.1; JOINED.
 CC PIR; B28816; KCHUM.
 CC PIR; S24324; S24324.
 CC PDB; 1MMP; 03-APR-96.
 CC PDB; 1MMQ; 03-APR-96.
 CC PDB; 1MMR; 03-APR-96.
 CC MEROPS; M10.008; -;
 CC MIM; I78990; -;
 CC InterPro: IPR000130; -;
 CC InterPro: IPR001818; -;
 CC Pfam: PF00413; Peptidase_M10; 1.
 CC PRINTS; PR00138; MATRININ.
 CC PROSITE; PS00142; ZINC_PROTEASE; 1.
 CC PROSITE; PS00546; CYSTEINE_SWITCH; 1.
 CC Hydrolyase; Metalloprotease; Zinc; zymogen; Calcium; 3D-structure;
 CC Collagen degradation; Extracellular matrix; Signal; Polymorphism.
 CC SIGNAL 1 17
 CC PROPEP 18 94
 CC CHAIN 95 267
 CC SITE 87 87
 CC METAL 214 214
 CC ACT_SITE 215 215
 CC METAL 218 218
 CC METAL 224 224
 CC VARIANT 77 77
 CC R -> H.
 CC
 CC SEQUENCE 267 AA; 29677 MW; F6BD1FC0ADA23603 CRC64;
 CC /FTIG=VAR_006729.
 CC
 CC Query Match 72.7%; Score 32; DB 1; Length 267;
 CC Best Local Similarity 83.3%; Pred. No. 31;
 CC Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 1 LMVPTY 6
 CC :|||||
 CC DB 231 VMVPTY 236
 CC
 CC RESULT 14
 CC MM07_RAT
 CC ID MM07_RAT
 CC AC P50280;
 CC
 CC STANDARD; PRT; 267 AA.
 CC
 CC Query Match 72.7%; Score 32; DB 1; Length 267;
 CC Best Local Similarity 83.3%; Pred. No. 31;
 CC Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 1 LMVPTY 6
 CC :|||||
 CC DB 231 VMVPTY 236
 CC
 CC RESULT 15
 CC PERE_ARATH
 CC ID PERE_ARATH
 CC AC P24102;
 CC DT 01-MAR-1992 (Rel. 21, Created)
 CC DT 01-MAR-1992 (Rel. 21, Last sequence update)
 CC DT 01-OCT-2000 (Rel. 40, Last annotation update)

DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE MATRILYSIN PRECURSOR (EC 3.4.24.23) (PUMP-1 PROTEASE) (UTERINE
 DE METALLOPROTEINASE) (MATRIX METALLOPROTEINASE-7) (MMP-7) (MATRIN).
 GN MMP7 OR MMP-7.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NCBI_TaxID=10116;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=uterus;
 RX MEDLINE=95332299; PubMed=7608162;
 RA Abramson S.R., Conner G.E., Negase H., Neuhaus I., Woessner J.F.;
 RT "Characterization of rat uterine matrilysin and its cDNA.
 RL J. Biol. Chem. 270:16016-16022(1995).
 CC -!- FUNCTION: DEGRADATES CASEIN, GELATINS OF TYPES I, III, IV, AND V,
 CC AND FIBRONECTIN. ACTIVATES PROCOLLAGENASE (BY SIMILARITY).
 CC -!- CATALYTIC ACTIVITY: CLEAVAGE OF 14-ALA-|-LEU-15 AND 16-TYR-|-
 CC LEU-17 IN B CHAIN OF INSULIN. NO ACTION ON COLLAGEN TYPES I, II,
 CC IV, V. CLEAVES GELATIN CHAIN ALPHA-2(I) > ALPHA-1(I).
 CC -!- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC
 CC METALLOPROTEASE) ALSO KNOWN AS MATRININ SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; L24374; AAA99432.1; -;
 CC HSSP; P09237; 1MMR.
 CC MEROPS; M10.008; -;
 CC InterPro: IPR000130; -;
 CC InterPro: IPR001818; -;
 CC Pfam: PF00413; Peptidase_M10; 1.
 CC PRINTS; PR00138; MATRININ.
 CC PROSITE; PS00142; ZINC_PROTEASE; 1.
 CC PROSITE; PS00546; CYSTEINE_SWITCH; 1.
 CC Hydrolyase; Metalloprotease; Zinc; zymogen; Calcium;
 CC Collagen degradation; Extracellular matrix; Signal.
 CC SIGNAL 1 20
 CC PROPEP 21 97
 CC CHAIN 98 267
 CC SITE 90 90
 CC METAL 217 217
 CC ACT_SITE 218 218
 CC METAL 221 221
 CC METAL 227 227
 CC SEQUENCE 267 AA; 29885 MW; EBA3C3D9527A4C7B CRC64;
 CC
 CC Query Match 72.7%; Score 32; DB 1; Length 267;
 CC Best Local Similarity 83.3%; Pred. No. 31;
 CC Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 1 LMVPTY 6
 CC :|||||
 CC DB 234 VMVPTY 239
 CC
 CC RESULT 15
 CC PERE_ARATH
 CC ID PERE_ARATH
 CC AC P24102;
 CC DT 01-MAR-1992 (Rel. 21, Created)
 CC DT 01-MAR-1992 (Rel. 21, Last sequence update)
 CC DT 01-OCT-2000 (Rel. 40, Last annotation update)

Job time: 306 sec

DE BASIC PEROXIDASE E PRECURSOR (EC 1.11.1.7).
GN PRXA.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91200671; PubMed=2016063;
RA Intapurk C., Higashimura N., Yamamoto K., Okada N., Shinmyo A.,
RA Takano M.;
RT "Nucleotide sequences of two genomic DNAs encoding peroxidase of
RL Arabidopsis thaliana";
CC Gene 98:237-241(1991).
CC -1- FUNCTION: REMOVAL OF H(2)O(2), OXIDATION OF TOXIC REDUCTANTS,
CC BIOSYNTHESIS AND DEGRADATION OF LIGNIN, DEFENSE RESPONSE TOWARD
CC WOUNDING AND METABOLISM OF AUXIN. THESE FUNCTIONS MIGHT BE
CC DEPENDENT ON EACH ISOZYME/ISOFORM IN EACH PLANT TISSUE.
CC -1- CATALYTIC ACTIVITY: DONOR + H(2)O(2) -> OXIDIZED DONOR + 2 H(2)O.
CC -1- COFACTOR: HEME.
CC -1- TISSUE SPECIFICITY: ROOTS.
CC -1- SIMILARITY: BELONGS TO THE PEROXIDASE FAMILY. PLANT PEROXIDASE
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M58381; AAA32842.1; -.
CC PIR; JU0458; JU0458.
CC HSP; P00433; IATJ
CC InterPro; IPR000823; -.
CC InterPro; IPR002016; -.
CC Pfam; PF00141; peroxidase; 1.
CC PRINTS; PR00458; PEROXIDASE.
CC PRINTS; PR00461; PLPEROXIDASE.
CC PROSITE; PS00435; PEROXIDASE_1; 1.
CC PROSITE; PS00436; PEROXIDASE_2; 1.
CC Oxidoreductase; Glycoprotein; Peroxidase; Heme; Multigene family;
KW Signal.
KW -----
FT SIGNAL 1 19
FT CHAIN 20 349 BASIC PEROXIDASE E.
FT ACT_SITE 67 67 BY SIMILARITY.
FT ACT_SITE 71 71 DISTAL HISTIDINE (BY SIMILARITY).
FT ACT_SITE 199 199 PROXIMAL HISTIDINE (HEME AXIAL LIGAND).
FT DISULFID 40 120 BY SIMILARITY.
FT DISULFID 73 78 BY SIMILARITY.
FT DISULFID 126 329 BY SIMILARITY.
FT DISULFID 206 238 BY SIMILARITY.
FT CARBOHYD 86 86 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 173 173 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 187 187 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 217 217 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 243 243 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 349 AA; 36172 MW; 6E02C5758C36AB8F CRC64;

Query Match 72.7%; Score 32; DB 1; Length 349;
Best Local Similarity 83.3%; Pred. No. 40;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYL 7

:|||||

Db 226 LYPTYL 231

Search completed: September 4, 2001, 15:54:09

Tue Sep 4 15:56:17 2001

us-09-630-345-1.rsp

Page 10

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2001, 15:48:47 ; Search time 80.44 seconds
(without alignments)
13.158 Million cell updates/sec

Title: US-09-630-345-1
Perfect score: 44
Sequence: 1 LMYPTYLK 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

al number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_16:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mnc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_unclassified:*
- 13: sp_vertebrate:*
- 14: sp_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35	79.5	378	10 Q9SHH5	Q9shh5 arabidopsis
2	35	79.5	483	13 Q98857	Q98857 cynops pyrr
3	35	79.5	1088	4 Q9H2T7	Q9h2t7 homo sapien
4	34	77.3	269	1 Q27159	Q27159 methanobact
5	34	77.3	291	5 Q94467	Q94467 dictyosteli
6	34	77.3	769	11 Q9ESN3	Q9esn3 mus musculu
7	33	75.0	238	14 Q9Q894	Q9qb94 yaba monkey
8	33	75.0	255	14 Q9DHM6	Q9dhm6 yaba-like d
9	33	75.0	261	4 Q9NRE1	Q9nrel homo sapien
10	33	75.0	261	4 Q9GZS2	Q9gzs2 homo sapien
11	33	75.0	261	4 Q9NR87	Q9nr87 homo sapien
12	33	75.0	262	2 Q9PIA0	Q9pia0 campylobact
13	33	75.0	294	10 Q9W432	Q9w432 populis tre
14	33	75.0	592	10 Q9S7Q5	Q9s7q5 arabidopsis
15	33	75.0	612	10 Q9SEJ5	Q9sej5 arabidopsis
16	33	75.0	1069	10 Q9LTZ8	Q9ltz8 arabidopsis
17	32	72.7	57	14 Q9WAG0	Q9wag0 mumps virus
18	32	72.7	187	3 Q13523	Q13523 saccharomyc
19	32	72.7	252	2 Q9L875	Q9l875 haemophilus

ALIGNMENTS

RESULT 1

Q9SHH5 PRELIMINARY; PRT; 378 AA.
ID Q9SHH5
AC Q9SHH5;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)
DE F20D23.14 PROTEIN.
GN F20D23.14
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, COLUMBIA;
RA Federspiel N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,
RA Altafi H., Araujo R., Huizar L., Rowley D., Buehler E., Dunn P.,
RA Gonzalez A., Kremenetskaia I., Kim C., Lenz C., Li J., Liu S.,
RA Luros S., Schwartz J., Shinn P., Toriumi M., Vysotskaia V.S.,
RA Walker M., Yu G., Ecker J., Theologis A., Davis R.W.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC007651; RAD50017.1; -;
DR HSSP; P05054; IRKD.
DR InterPro; IPR002139; -;
DR InterPro; IPR002173; -;
DR Pfam; PF00294; pfkB; 1.
DR PRINTS; PR00990; RIBOKINASE.
DR PROSITE; PS00584; PFYB_KINASES_2; 1.
SQ SEQUENCE 378 AA; 40058 MW; 7F7D93E71A13AD54 CRC64;

Query Match 79.5%; Score 35; DB 10; Length 378;

Best Local Similarity 100.0%; Pred. No. 50;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMYPTY 6

Db 117 LMYPTY 122

RESULT 2
Q98857 ID Q98857 PRELIMINARY; PRT; 483 AA.
AC Q98857
DT 01-FEB-1997 (TReMBLrel. 02, Created)
DT 01-FEB-1997 (TReMBLrel. 02, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE STROMELYSIN-1/2-A.
OS Cynops pyrrhogaster (Japanese common newt).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Caudata; Salamandroidae; Salamandridae; Cynops.
OX NCBI_TaxID=8330;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIMB;
RX MEDLINE=96270627; PubMed=8692902;
RA Miyazaki K., Uchiyama K., Imokawa Y., Yoshizato K.;
RT "Cloning and characterization of cDNAs for matrix metalloproteinases
of regenerating newt limbs.";
Proc. Natl. Acad. Sci. U.S.A. 93:6819-6824(1996).
EMBL; D82053; BAA11524.1;
HSSP; P08254; IUSN.
DR MEROPS; M10.018;
DR InterPro; IPR000130;
DR InterPro; IPR000585;
DR InterPro; IPR001818;
DR Pfam; PF00045; hemopexin; 4.
DR PRINTS; PR00138; Peptidase M10; 1.
DR PRIMATE; PR00138; MATRIXIN.
DR PROSITE; PS00024; HEMOPEXIN; UNKNOWN_1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
DR SMART; SM00120; HX; 1.
SQ SEQUENCE 483 AA; 54171 MW; 4ABDF98A36E90EAF CRC64;

Query Match 79.5%; Score 35; DB 13; Length 483;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LMPTY 6
DB 233 LMPTY 238
Q98857 ID Q98857 PRELIMINARY; PRT; 1088 AA.
AC Q98857
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DE RANBP17
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthalia; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20568271; PubMed=11024021;
RA Kutay U., Hartmann E., Teichner N., Calado A., Carmo-Fonseca M.,
RA Prehn S., Kraft R., Gorlich D., Bischoff F.R.;
RT "Identification of two Novel RanGTP-Binding Proteins Belonging to the
Importin beta Superfamily".
J. Biol. Chem. 275:40163-40168(2000).
DR EMBL; AF222747; AAG4255.1;
SQ SEQUENCE 1088 AA; 124374 MW; A95D599388EAEFB CRC64;

Query Match 79.5%; Score 35; DB 13; Length 483;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LMPTY 6
DB 233 LMPTY 238

RESULT 3
Q98277 ID Q98277 PRELIMINARY; PRT; 1088 AA.
AC Q98277
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DE RANBP17
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthalia; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20568271; PubMed=11024021;
RA Kutay U., Hartmann E., Teichner N., Calado A., Carmo-Fonseca M.,
RA Prehn S., Kraft R., Gorlich D., Bischoff F.R.;
RT "Identification of two Novel RanGTP-Binding Proteins Belonging to the
Importin beta Superfamily".
J. Biol. Chem. 275:40163-40168(2000).
DR EMBL; AF222747; AAG4255.1;
SQ SEQUENCE 1088 AA; 124374 MW; A95D599388EAEFB CRC64;

Query Match 79.5%; Score 35; DB 4; Length 1088;
Best Local Similarity 100.0%; Pred. No. 1,5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 MYPTYL 7

Db 744 MYPTYL 749
RESULT 4
Q27159 ID Q27159 PRELIMINARY; PRT; 269 AA.
AC Q27159
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE HYPOTHETICAL 31.5 KDA PROTEIN.
GN MTH1087.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
OC Methanothermobacter.
OX NCBI_TaxID=145262;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DELTA H;
RX MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., DuBois J.,
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lamm W., Pothier B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jiواني N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McDougall S., Shimer G., Goyal A., Picirovski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Nolling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
deltaH: functional analysis and comparative genomics.";
J. Bacteriol. 179:7135-7155(1997).
RL EMBL; AE000879; AAB85576.1;
KW Hypothetical protein.
SQ SEQUENCE 269 AA; 31513 MW; 4B7D64BDC6B8EA7 CRC64;

Query Match 77.3%; Score 34; DB 1; Length 259;
Best Local Similarity 62.5%; Pred. No. 55;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 LMPTYLK 8
DB 173 MYPTYLK 180

RESULT 5
Q94467 ID Q94467 PRELIMINARY; PRT; 291 AA.
AC Q94467
DT 01-FEB-1997 (TReMBLrel. 02, Created)
DT 01-FEB-1997 (TReMBLrel. 02, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE SSRNA-BINDING PROTEIN.
OS Dictyostelium discoideum (slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX2
RX MEDLINE=98086164; PubMed=9426195;
RA Fucosier P., Nellen W.;
RT "Functional activity and developmental regulation of DdRBP1, a RNA
binding protein in Dictyostelium discoideum.";
Biol. Chem. 378:1353-1360(1997).
DR EMBL; Y07952; XA69250.1;
DR HSSP; P09651; IPR00050;
DR InterPro; IPR000504;
DR Pfam; PF00076; RNP_1;
DR PROSITE; PS00030; RNP_1; UNKNOWN_1.
DR SMART; SM00360; RNP_1;
SQ SEQUENCE 291 AA; 31767 MW; D8B717C82AE0C2D5 CRC64;


```
Query Match 77.3%; Score 34; DB 5; Length 291;
Best Local Similarity 75.0%; Pred. No. 60;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LMYPYLYK 8
DB 221 LKYPTEFK 228
      | |||:|
      | |||:|

RESULT 6
Q9ESN3 PRELIMINARY; PRT; 769 AA.
AC Q9ESN3;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DE M83 PROTEIN.
DE M83.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=THYMUS;
RX PubMed=11006113;
RA Motohashi T., Miyoshi S., Osawa M., Eyre H.J., Sutherland G.R.,
RA Matsuda Y., Nakamura Y., Shibuya A., Iwama A., Nakauchi H.;
RT "Molecular cloning and chromosomal mapping of a novel five-span
RT transmembrane protein gene, M83.";
RL Biochem. Biophys. Res. Commun. 276:244-250(2000).
DR EMBL; AB045293; BAB16377.1; -.
SQ SEQUENCE 769 AA; 85328 MW; A0B886EE571B7B60 CRC64;

Query Match 77.3%; Score 34; DB 11; Length 769;
Best Local Similarity 75.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LMYPYLYK 8
DB 203 LSYPSYLYK 210
      | |||:|
      | |||:|

RESULT 7
Q9QB94 PRELIMINARY; PRT; 238 AA.
AC Q9QB94;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DE YB-B4R PROTEIN.
DE YB-B4R.
OS Yaba monkey tumor virus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Yatapoxvirus
NCBI_TaxID=38804;
RN [1]
RP SEQUENCE FROM N.A.
RA Amano H., Morikawa S., Ueda Y., Miyamura T.;
RT "Nucleotide sequence of the central 50kbp region of Yaba virus DNA.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95248279; PubMed=7730796;
RA Amano H., Ueda Y., Miyamura T.;
RT "Identification and characterization of the thymidine kinase gene of
RT yaba virus.";
RL J. Gen. Virol. 76:1109-1115(1995).
DR EMBL; AB015885; BAA8801.1; -.
DR InterPro; IPR000086; -.
DR Pfam; PF00293; mult. 1.
DR PROSITE; PS00893; MUT; 1.

SQ SEQUENCE 238 AA; 27721 MW; 393119A65D04DA82 CRC64;

Query Match 75.0%; Score 33; DB 14; Length 238;
Best Local Similarity 75.0%; Pred. No. 75;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LMYPYLYK 8
DB 89 LKYPNYLK 96
      | |||:|
      | |||:|

RESULT 8
Q9DHM6 PRELIMINARY; PRT; 255 AA.
AC Q9DHM6;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DE 87R PROTEIN.
DE 87R.
OS Yaba-like disease virus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Yatapoxvirus.
NCBI_TaxID=132475;
RN [1]
RP SEQUENCE FROM N.A.
RA Lee H.J., Smith G.L.;
RT "The genome sequence of yaba-like disease virus, a yatapoxvirus.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Lee H.J.;
RL Thesis (2000), Sir William Dunn School of Pathology, University of
DR EMBL; AJ293568; CAC21325.1; -.
SQ SEQUENCE 255 AA; 29981 MW; DAB3F6002DCFE5CD CRC64;

Query Match 75.0%; Score 33; DB 14; Length 255;
Best Local Similarity 75.0%; Pred. No. 81;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LMYPYLYK 8
DB 89 LKYPNYLK 96
      | |||:|
      | |||:|

RESULT 9
Q9NRE1 PRELIMINARY; PRT; 261 AA.
AC Q9NRE1;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DE 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE MATRIX METALLOPROTEINASE MMP-26.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20285359; PubMed=10824119;
RA Benoit De Coignac A., Elson G., Delneste Y., Magistrelli G.,
RA Jeannin P., Aubry J.P., Berthier O., Schmitt D., Bonnefoy J.Y.,
RA Gauchat J.F.;
RT "Cloning of MMP-26 A novel matrilysin-like proteinase.";
RL Eur. J. Biochem. 267:3323-3329(2000).
DR EMBL; AF20354; AAF80180.1; -.
DR InterPro; IPR000130; -.
DR Pfam; PF00413; Peptidase M10; 1.
DR PRINTS; PR00138; MATRIXIN.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
```

BN	[1]
RP	SEQUENCE FROM N.A.
RA	Guria J.A., Lopez-Otin C.
RH	Human matrilysin-2: Identification and functional characterization."
RI	Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
RN	[2]
PP	SEQUENCE FROM N.A.
PA	Sarchenko G.N., Strongin A.Y.;
PL	"Matrix metalloproteinase-26 gene.";
PT	Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
RL	EMBL; AJ251531; CAC08821.1; -
DR	EMBL; AF291665; AAG02470.1; -
DQ	SEQUENCE 261 AA; 29708 MW; A66D0DBE3ED7DE76 CRC64;
SD	
Query Match	75.0%; Score 33; DB 4; Length 261;
Best Local Similarity	83.3%; Pred. No. 83;
Matches	5; Conservative 1; Mismatches 0; Indels 0; Gaps
QY	1 LMPYTYL 6
Db	-
	225 INPTY 230
RESULT	12
ID	O9PIAO PRELIMINARY; PRT; 262 AA.
AC	O9PIAO;
DT	01-OCT-2000 (TEMBREL_15, Created)
DT	01-OCT-2000 (T-EMBEL_15, Last sequence update)
DT	01-MAR-2001 (T-EMBEL_16, Last annotation update)
DE	SHKIMATE 5-DEHYDROGENASE (EC 1.1.1.25).
GN	ARO OR CJ0405.
GS	Campylobacter jejuni.
OC	Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
CC	Campylobacter.
OX	NCBI_TaxID=197;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=NCTC 11168;
RX	MEDLINE=20150912; PubMed=10688204;
RA	Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,
RA	Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,
RA	Jagels K., Karyshev A.V., Moule S., Pallen M.J., Penn C.W.,
RA	Quail M.A., Rajandream A.G., Rutherford K.M., Van Vleet A.H.M.,
RA	Whitehead S., Barrell B.G.;
RT	"The genome sequence of the food-borne pathogen Campylobacter jejuni
RT	reveals hypervariable sequences.";
SL	Nature 403:665-668(2000).
DL	EMBL; ALI39075; CAB74241.1; -
DR	InterPro; IPR002907; -
DD	Tran; PF01488; Shikimate_DH; 1.
DR	SEQUENCE 262 AA; 29977 MW; 22FABC37B66BDFC2 CRC64;
SD	
Query Match	75.0%; Score 33; DB 2; Length 262;
Best Local Similarity	85.7%; Pred. No. 84;
Matches	6; Conservative 0; Mismatches 1; Indels 0; Gaps
QY	1 LMPYTYL 7
Db	-
	153 LAYPYTL 159
RESULT	13
ID	O9M432 PRELIMINARY; PRT; 294 AA.
AC	O9M432;
DT	01-OCT-2000 (TEMBREL_15, Created)
DT	01-OCT-2000 (T-EMBEL_15, Last sequence update)
DT	01-MAR-2001 (T-EMBEL_16, Last annotation update)
DE	ALTERNATIVE OXIDASE.
GN	AOXIB.

OS Populus tremula x Populus tremuloides.
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
 OC Malpighiales; Salicaceae; Populus.
 OX NCBI_TaxID=47664;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LEAF;
 RA Hasenfratz-Sauder M.P., Jolivet Y., Gerard J., Dizengremel P.;
 RT "Changes in alternative oxidase expression during ozone fumigation of
 RT poplar trees.";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ271889; CAB72441.1; -
 DR InterPro: IPR002680; -
 DR Pfam: PF01786; AOX; 1.
 SQ SEQUENCE 294 AA; 33665 MW; DA7EFAA4A5D8B65A CRC64;

Query Match 75.0%; Score 33; DB 10; Length 294;

Best Local Similarity 85.7%; Pred. No. 94;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LMYPTYLK 7
 | | | | |
 DB 103 LKYPTYL 109

RESULT 14

Q9S7Q5 PRELIMINARY; PRT; 592 AA.
 AC Q9S7Q5;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DE FLC9.34 PROTEIN.
 GN FLC9.34 OR F28J7.21.
 OS Arabidopsis thaliana (Mouse-ear cross).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.,
 RA Ronning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
 RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;
 RT "Arabidopsis thaliana chromosome III BAC FLC9 genomic sequence.";
 RT Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC011664; AAF14847.1; -
 DR EMBL: AC010797; AAF03440.1; -
 SQ SEQUENCE 592 AA; 66531 MW; 65BACALIC886D43D0 CRC64;

Query Match 75.0%; Score 33; DB 10; Length 592;

Best Local Similarity 75.0%; Pred. No. 2e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LMYPTYLK 8
 | | | | |
 DB 3 LLYETYLK 10

RESULT 15

Q9SEJ5 PRELIMINARY; PRT; 612 AA.
 AC Q9SEJ5;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
 DE PHOSPHOGLUCOSE ISOMERASE PRECURSOR.
 GN PGI.
 OS Arabidopsis thaliana (Mouse-ear cross).

OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yu T.S., Lue W.L., Wang S.M., Chen J.;
 RT "Mutation of Arabidopsis chloroplastic phosphoglucose isomerase
 RT affects starch synthesis and floral initiation.";
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF120494; AAF24124.1; -
 DR InterPro: IPR001672; -
 DR Pfam: PF00342; PGI; 2.
 DR PRINTS: PR00662; G6PISOMERASE.
 KW Transit peptide; Isomerase.
 FT TRANSIT 1 73 POTENTIAL.
 SQ SEQUENCE 612 AA; 67149 MW; 42D38E445816AC24 CRC64;

Query Match 75.0%; Score 33; DB 10; Length 612;

Best Local Similarity 62.5%; Pred. No. 2.1e+02;
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMYPTYLK 8
 | | | | |
 DB 33 LFFPTYLQ 40

Search completed: September 4, 2001, 15:53:35
 Job time: 288 sec

Tue Sep 4 15:56:18 2001

us-09-630-345-1.rspt

Page 6

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2001, 15:50:25 ; Search time 77.49 Seconds
(without alignments)
6.259 Million cell updates/sec

Title: US-09-630-345-2

Perfect score: 41

Sequence: 1 XMYPTYLK 8

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

al number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_0601.*

- 1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.*
- 2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.*
- 3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.*
- 4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.*
- 5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.*
- 6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.*
- 7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.*
- 8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.*
- 9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.*
- 10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.*
- 11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.*
- 12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.*
- 13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.*
- 14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.*
- 15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.*
- 16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.*
- 17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.*
- 18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.*
- 19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.*
- 20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.*
- 21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.*
- 22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred, No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	97.6	8	19 AAW50937	Vasoactive intesti
2	40	97.6	8	21 AAB08298	Vasoactive intesti
3	40	97.6	8	21 AAB08306	Amino acid sequenc
4	35	85.4	8	21 AAB08304	Amino acid sequenc
5	35	85.4	8	21 AAB08305	Amino acid sequenc
6	35	85.4	450	21 AAB63104	Human secreted pro
7	32	78.0	807	21 AAY57288	Human GPCR protein
8	32	78.0	986	21 AAY40440	Human brain-deri
9	31	75.6	8	20 AAW75457	Mammalian tub prot
10	31	75.6	12	20 AAW97073	Peptidomimetic cap
11	31	75.6	58	20 AAY12222	Human 5' EST secre

12	31	75.6	173	22	AAW74618	Human 72 kDa gelat
13	31	75.6	173	22	AAW68682	Matrilysin catalyt
14	31	75.6	261	19	AAW76253	Human matrilysin-1
15	31	75.6	264	19	AAW76254	Human matrilysin p
16	31	75.6	271	17	AAW96211	Human recombinant
17	31	75.6	297	21	AAW43585	Human cancer assoc
18	31	75.6	325	14	AAW42750	ced-4 gene II prod
19	31	75.6	459	18	AAW36485	Mouse tub form I.
20	31	75.6	459	21	AAW26400	Mouse tub form I p
21	31	75.6	460	18	AAW36488	Human TUB Form 6.
22	31	75.6	460	21	AAW26903	Human TUB Form 6 p
23	31	75.6	462	22	AAW49983	Murine macrophage
24	31	75.6	469	21	AAW10655	BPV1 L1 fusion pro
25	31	75.6	469	21	AAW10660	BPV2 L1 fusion pro
26	31	75.6	483	20	AAW93163	Human MMP-20 prote
27	31	75.6	495	6	AAW51099	Sequence of bovine
28	31	75.6	495	20	AAW88482	Bovine papillomavi
29	31	75.6	505	18	AAW36486	Mouse TUB Form II.
30	31	75.6	505	18	AAW10728	Mouse tub gene pro
31	31	75.6	505	19	AAW54367	Mouse tub polypept
32	31	75.6	505	20	AAW75450	Mouse wild type tu
33	31	75.6	505	21	AAW26901	Mouse tub form II
34	31	75.6	506	18	AAW36497	Human TUB Form 4.
35	31	75.6	506	18	AAW10729	Human tub gene pro
36	31	75.6	506	19	AAW54368	Human tub polypept
37	31	75.6	506	20	AAW75451	Human wild type tu
38	31	75.6	506	21	AAW26910	Human TUB form 4 5
39	31	75.6	512	18	AAW36496	Human TUB Form 3.
40	31	75.6	512	21	AAW26912	Human TUB form 3 5
41	31	75.6	518	18	AAW36495	Human TUB Form 2.
42	31	75.6	518	21	AAW26909	Human TUB form 2 5
43	31	75.6	561	18	AAW36489	Human TUB Form 1.
44	31	75.6	561	21	AAW26904	Human TUB Form 1 p
45	30	73.2	73	14	AAW40041	Filistata peptide

ALIGNMENTS

RESULT 1

AAW50937

ID AAW50937 standard; peptide; 8 AA.

XX AC AAW50937;

XX DT 31-JUL-1998 (first entry)

XX DE Vasoactive intestinal peptide receptor binding inhibitor (VIP2).

XX KW Vasoactive intestinal peptide; VIP; antagonist; somatostatin; bombesin;

XX KW Substance P; cancer; inhibition.

XX OS Synthetic.

XX PN EP835662-A2.

XX PD 15-APR-1998.

XX PF 11-DEC-1996; 96EP-0309012.

XX PR 08-OCT-1996; 96US-0727679.

XX PR 16-AUG-1996; 96IN-0001822.

XX (NAIM-) NAT INST IMMUNOLOGY.

XX Jaggi M, Mukherjee R;

XX WPI; 1998-208959/19.

XX Composition containing analogues of vasoactive intestinal peptide, somatostatin - bombesin and substance P, for treatment of tumours and for inhibiting over-expression of these peptide(s)

PS Claim 1; Page 4; 49pp; English.

XX The invention relates to a new composition which comprises: (i) the
CC somatostatin analogue SOM2 AGCKNFFDKPTSDC (3-14 disulphide bridge),
CC and (ii) at least 4 of the peptides: antagonist of vasoactive
CC intestinal peptide (VIP1); VIP receptor-binding inhibitor (VIP2); VIP
CC receptor antagonist (VIP3); somatostatin analogue (SOM1); bombesin
CC antagonist (BOM1) and substance P antagonist (SPI). Also claimed are
CC more general compositions containing peptide analogues of somatostatin,
CC VIP, bombesin and substance P. The compositions are used in human or
CC veterinary medicine: (a) to kill (or inhibit multiplication of) tumour
CC or cancer cells, particularly for treatment of leukaemia, lymphoma,
CC adenocarcinoma of stomach, pancreas or prostate, or cancer of lung,
CC breast, kidney or particularly rectum and colon, and (b) to prevent,
CC inhibit or modulate over-expression of, e.g. VIP. A wide range of cancer
CC cells express receptors for VIP, somatostatin, bombesin and/or substance
CC P. The present sequence represents VIP receptor-binding inhibitor
CC (VIP2).

XX Sequence 8 AA;

Query Match 97.6%; Score 40; DB 19; Length 8;
Best Local Similarity 100.0%; Pred. No. 3.4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
DB 2 myptylk 8

RESULT 2
AAB08298
ID AAB08298 standard; peptide; 8 AA.

XX AAB08298;
XX 04-DEC-2000 (first entry)
XX Vasoactive intestinal peptide (VIP) analogue VIP2.
XX Vasoactive intestinal peptide; VIP; analogue; somatostatin; SOM1; SOM2;
KW VIP1; VIP2; VIP3; BOM1; bombesin; SPI; substance P; MuJ-7; tumour growth;
KW tumour angiogenesis; metastasis; cancer; angiogenesis; adenocarcinoma;
KW leukaemia; lymphoma.

OS Synthetic.

WO200047221-A1.

17-AUG-2000.

11-FEB-2000; 2000WO-US03559.

11-FEB-1999; 99US-0248381.

(NAIN-) NAT INST IMMUNOLOGY.

(DABU-) DABUR RES FOUND.

(CORD/) CORD J I.

Mukherjee R, Jaggi M, Prasad S, Burman AC, Rajendran P, Mathur A;
PI Singh At.

WPI; 2000-549083/50.

XX Novel therapeutically active composition comprising at least 5
FT peptides, useful for treating angiogenesis especially as a result of
PT adenocarcinomas -

XX Disclosure; Page 8; 42pp; English.

PS The present sequence represents an analogue of vasoactive intestinal
CC peptide (VIP). The specification describes therapeutically active

CC compositions comprising at least one analogue of somatostatin (chosen
CC from SOM1 and SOM2), and at least four analogues chosen from VIP1 (a
CC VIP antagonist), VIP2 (a VIP receptor binding inhibitor), VIP3 (a VIP
CC receptor antagonist), BOM1 (a bombesin antagonist), and SPI (a substance
CC P antagonist). The combination of these 7 analogues is known as MuJ-7.
CC MuJ-7 is used as an anticancer drug to restrict tumour growth and spread
CC by inhibiting tumour angiogenesis. MuJ-7, in addition, inhibits
CC metastasis through its antiangiogenic activity in all cancers. The
CC peptides are useful for the treatment and prevention of angiogenesis,
CC especially as a result of adenocarcinomas of the colon, breast, lung,
CC prostate, kidney, leukemias or lymphomas.

XX Sequence 8 AA;

Query Match 97.6%; Score 40; DB 21; Length 8;
Best Local Similarity 100.0%; Pred. No. 3.4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
DB 2 myptylk 8

RESULT 3
AAB08306
ID AAB08306 standard; peptide; 8 AA.

XX AAB08306;

XX 04-DEC-2000 (first entry)

XX Amino acid sequence of an antiangiogenic peptide.

XX Vasoactive intestinal peptide; VIP; analogue; somatostatin; SOM1; SOM2;
KW VIP1; VIP2; VIP3; BOM1; bombesin; SPI; substance P; MuJ-7; tumour growth;
KW tumour angiogenesis; metastasis; cancer; angiogenesis; adenocarcinoma;
KW leukaemia; lymphoma.

OS Synthetic.

Key Location/Qualifiers

FT Misc-difference 1 /note= "optionally D-form residue"

FT Misc-difference 3 /note= "optionally D-form residue"

FT Misc-difference 6 /note= "optionally D-form residue"

FT Misc-difference 7 /note= "optionally D-form residue"

FT Misc-difference 7 /note= "optionally D-form residue"

XX WO200047221-A1.

17-AUG-2000.

11-FEB-2000; 2000WO-US03559.

11-FEB-1999; 99US-0248381.

(NAIN-) NAT INST IMMUNOLOGY.

(DABU-) DABUR RES FOUND.

(CORD/) CORD J I.

Mukherjee R, Jaggi M, Prasad S, Burman AC, Rajendran P, Mathur A;
PI Singh At.

WPI; 2000-549083/50.

XX Novel therapeutically active composition comprising at least 5
FT peptides, useful for treating angiogenesis especially as a result of
PT adenocarcinomas -

XX Claim 11; Page 31; 42pp; English.

XX AAB08304-15 represent peptides which have an antiangiogenic effect. The
 CC specification describes therapeutically active compositions comprising
 CC at least one analogue of somatostatin (chosen from SOM1 and SOM2), and
 CC at least four analogues chosen from vasoactive intestinal peptide (VIP)
 CC 1 (a VIP antagonist), VIP2 (a VIP receptor binding inhibitor), VIP3 (a
 CC VIP receptor antagonist), BOM1 (a bombesin antagonist), and Sp1 (a
 CC substance P antagonist). The combination of these 7 analogues is known as
 CC MuJ-7. MuJ-7 is used as an anticancer drug to restrict tumour growth and
 CC spread by inhibiting tumour angiogenesis. MuJ-7, in addition, inhibits
 CC metastasis through its antiangiogenic activity in all cancers. The
 CC peptides are useful for the treatment and prevention of angiogenesis,
 CC especially as a result of adenocarcinomas of the colon, breast, lung,
 CC prostate, kidney, leukemias or lymphomas.
 XX Sequence 8 AA;
 SQ

Query Match 97.6%; Score 40; DB 21; Length 8;
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 MYPTYLK 8
 |||||
 Db 2 myptylk 8

RESULT 4
 AAB08304
 ID AAB08304 standard; peptide; 8 AA.
 XX AC AAB08304;
 XX DT 04-DEC-2000 (first entry)
 XX DE Amino acid sequence of an antiangiogenic peptide.
 XX Vasoactive intestinal peptide; VIP; analogue; somatostatin; SOM1; SOM2;
 KW VIP1; VIP2; VIP3; BOM1; bombesin; Sp1; substance P; MuJ-7; tumour growth;
 KW tumour angiogenesis; metastasis; cancer; angiogenesis; adenocarcinoma;
 KW leukaemia; lymphoma.
 OS Synthetic.
 XX FH Key Location/Qualifiers
 FT Modified-site 1 /label= Aib
 FT /note= "alpha-aminoisobutyric acid"
 FT Modified-site 7 /label= Aib, Deg
 FT /note= "alpha-aminoisobutyric acid or alpha, alpha
 diethyl glycine"
 FT WO200047221-A1.
 PN 17-AUG-2000.
 XX 11-FEB-2000; 2000WO-US03559.
 XX 11-FEB-1999; 99US-0248381.
 XX (NAIM-) NAT INST IMMUNOLOGY.
 PA (DABU-) DABUR RES FOUND.
 PA (CORD/) CORD J I.
 XX Mukherjee R, Jaggi M, Prasad S, Burman AC, Rajendran P, Mathur A;
 PI Singh AT;
 XX WPI; 2000-549083/50.
 XX Novel therapeutically active composition comprising at least 5
 PT peptides, useful for treating angiogenesis especially as a result of
 PT adenocarcinomas -

XX Claim 18; Page 36; 42pp; English.
 XX AAB08304-15 represent peptides which have an antiangiogenic effect. The
 CC specification describes therapeutically active compositions comprising
 CC at least one analogue of somatostatin (chosen from SOM1 and SOM2), and
 CC at least four analogues chosen from vasoactive intestinal peptide (VIP)
 CC 1 (a VIP antagonist), VIP2 (a VIP receptor binding inhibitor), VIP3 (a
 CC VIP receptor antagonist), BOM1 (a bombesin antagonist), and Sp1 (a
 CC substance P antagonist). The combination of these 7 analogues is known as
 CC MuJ-7. MuJ-7 is used as an anticancer drug to restrict tumour growth and
 CC spread by inhibiting tumour angiogenesis. MuJ-7, in addition, inhibits
 CC metastasis through its antiangiogenic activity in all cancers. The
 CC peptides are useful for the treatment and prevention of angiogenesis,
 CC especially as a result of adenocarcinomas of the colon, breast, lung,
 CC prostate, kidney, leukemias or lymphomas.
 XX Sequence 8 AA;
 SQ

Query Match 85.4%; Score 35; DB 21; Length 8;
 Best Local Similarity 85.7%; Pred. No. 3.4e+05;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 MYPTYLK 8
 |||||
 Db 2 myptylk 8

RESULT 5
 AAB08305
 ID AAB08305 standard; peptide; 8 AA.
 XX AC AAB08305;
 XX DT 04-DEC-2000 (first entry)
 XX DE Amino acid sequence of an antiangiogenic peptide.
 XX Vasoactive intestinal peptide; VIP; analogue; somatostatin; SOM1; SOM2;
 KW VIP1; VIP2; VIP3; BOM1; bombesin; Sp1; substance P; MuJ-7; tumour growth;
 KW tumour angiogenesis; metastasis; cancer; angiogenesis; adenocarcinoma;
 KW leukaemia; lymphoma.
 OS Synthetic.
 XX FH Key Location/Qualifiers
 FT Misc-difference 1 /note= "D-form residue"
 FT Modified-site 7 /label= Aib, Deg
 FT /note= "alpha-aminoisobutyric acid or alpha, alpha
 diethyl glycine"
 FT WO200047221-A1.
 PN 17-AUG-2000.
 XX 11-FEB-2000; 2000WO-US03559.
 XX 11-FEB-1999; 99US-0248381.
 XX (NAIM-) NAT INST IMMUNOLOGY.
 PA (DABU-) DABUR RES FOUND.
 PA (CORD/) CORD J I.
 XX Mukherjee R, Jaggi M, Prasad S, Burman AC, Rajendran P, Mathur A;
 PI Singh AT;
 XX WPI; 2000-549083/50.
 XX Novel therapeutically active composition comprising at least 5
 PT peptides, useful for treating angiogenesis especially as a result of

“

PT Human G protein coupled protein receptor peptides useful for the
 PT prevention, diagnosis and treatment of cell proliferative, neurological
 PT and immune disorders -

XX
 PS Claim 1; Page 65-67; 71pp; English.

XX
 CC The invention provides human G protein coupled protein receptor (HGPRP)
 CC polypeptides and polynucleotides encoding them. The polypeptides can be
 CC produced by standard recombinant methodology. The polynucleotides and
 CC polypeptides may be used in the prevention, treatment and diagnosis of
 CC diseases associated with their inappropriate expression. Diseases that
 CC can be treated are cell proliferative disorders (e.g. arteriosclerosis,
 CC atherosclerosis and hepatitis), cancers (e.g. leukemia, melanomas and
 CC adenocarcinoma), immune disorders (e.g. anemia, asthma and Crohn's
 CC disease) and neurological disorders (e.g. epilepsy, Alzheimer's disease
 CC and Parkinson's disease). The anti-HGPRP antibodies may also be used as
 CC diagnostic agents for detecting the presence of HGPRP polypeptides in
 CC samples (e.g. by enzyme linked immunosorbent assay (ELISA)). Sequences
 CC AAY57283-288 represent the HGPRP polypeptides.

XX Sequence 807 AA;

Query Match 78.0%; Score 32; DB 21; Length 807;
 Best Local Similarity 85.7%; Pred. No. 2.9e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
 | | | | |
 Db 205 mlptylk 211

RESULT 8

AAAY40440
 ID AAY40440 standard; Protein; 986 AA.

AC AAY40440;

DT 26-MAY-2000 (first entry)

XX Human brain-derived G-protein coupled receptor protein.

XX G-protein coupled receptor protein; human; brain; gene therapy;
 KW genetic disease; screening assay.

XX Homo sapiens.

XX WO200008053-A1.

PN 17-FEB-2000.

PF 05-AUG-1999; 99WO-JP04233.

XX 07-AUG-1998; 98JP-0225059.

XX (TAKE) TAKEDA CHEM IND LTD.
 PA (KAZU-) KAZUSA DNA RES INST.

XX Ohara O, Nagase T, Nomura N, Mogi S, Yamamoto K, Kurokawa T;

PI WPI; 2000-195555/17.

DR N-PSDB; AA287930.

XX Human brain-derived G-protein coupled receptor protein and encoding DNA
 PT used in gene therapy, is also useful for e.g. identifying ligands and
 PT raising antibodies and antisera -

XX Claim 1; Fig 1; 80pp; Japanese.

XX This represents a human brain-derived G-protein coupled receptor
 CC protein. The polypeptide can be expressed by standard recombinant
 CC methodology. The novel G-protein coupled receptor protein can be used
 CC for identifying ligands; raising antibodies and antisera; developing

CC receptor-binding assay system; and screening for drug candidates. The
 CC encoding polynucleotide can be used in; probing for diagnostic genes;
 CC constructing PCR primers; making transgenic animals; and in gene therapy.
 XX
 SQ Sequence 986 AA;

Query Match 78.0%; Score 32; DB 21; Length 986;
 Best Local Similarity 85.7%; Pred. No. 3.6e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
 | | | | |
 Db 384 mlptylk 390

RESULT 9

AAW75457

ID AAW75457 standard; peptide; 8 AA.

XX AAW75457;

XX 27-APR-1999 (first entry)

XX Mammalian tub protein tyrosine phosphorylation site.

XX Human; wild type; tubby; identification; SH2 domain; mammal; obesity;
 KW body weight disorder; cachexia; anorexia.

XX Homo sapiens.

XX Mus sp.

XX US5861239-A.

XX 19-JAN-1999.

XX 02-SEP-1997; 97US-0922267.

XX 02-SEP-1997; 97US-0922267.

PR 12-APR-1996; 96US-0631200.

PR 28-MAR-1997; 97US-0829553.

XX (MILL-) MILLENNIUM PHARM INC.

XX Kapeller R, Kiehn PW, Moore KJ;

XX WPI; 1999-130383/11.

XX Identifying compounds which modulate tub protein activity - by
 PT detecting compounds which alter the interaction of tub protein with
 PT a SH2 containing peptide, used to develop agents for treating e.g.
 PT obesity, cachexia or anorexia

XX Disclosure; Column 13; 95pp; English.

XX This sequence represents a tyrosine phosphorylation site/SH2 docking
 CC domain found in the amino acid sequence of the mouse and human "tub"
 CC proteins (AAW75450 and AAW75451 respectively). The invention relates to
 CC a method for identifying compounds that modulate tub protein activity,
 CC especially its interaction with proteins containing an SH2 domain. The
 CC method can be used for identifying compounds which modulate tub protein
 CC activity for use in the treatment of mammalian body weight disorders
 CC including obesity, cachexia and anorexia.

XX Sequence 8 AA;

Query Match 75.6%; Score 31; DB 20; Length 8;
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
 | | | | |

Db 2 mypty 6

RESULT 10

AAW97073

ID AAW97073 standard; peptide; 12 AA.

XX

AC AAW97073;

XX

DT 29-APR-1999 (first entry)

XX

DE Peptidomimetic capable of inhabiting CD28 and/or CTLA-4 interactions.

XX

XX Peptidomimetic; CD28; CTLA-4; CD80; CD86; B7-1; B7-2; diabetes;

KW

immune system disease; autoimmune disease; psoriasis; multiple sclerosis;

KW

lupus erythematosus; rheumatoid arthritis; transplant rejection; cancer.

XX

OS Synthetic.

XX

WO9856401-A1.

PD

17-DEC-1998.

XX

PF 11-JUN-1998; 98WO-US12312.

XX

PR 12-JUN-1997; 97US-0049470.

XX

PA (ISTF) ARS APPLIED RES SYSTEMS HOLDING NV.

XX

PI Blechner S, El Tayar N, Jameson B, Tepper M;

XX

WPI; 1999-080854/07.

XX

New peptidomimetic compounds - useful in the prophylaxis, diagnosis and treatment of pathologies and disorders, which are improved by inhibition of CD28 and/or CTLA-4 interaction with CD80 (B7-1) and CD86 (B7-2)

XX

PS Disclosure; Page 9; 62pp; English.

XX

AAW97053-82 represent peptidomimetic compounds that inhibit CD28 and/or CTLA-4 interactions with CD80 and CD86. The peptides are used in pharmaceutical compositions for the prophylaxis, diagnosis and treatment of pathologies and disorders, which are improved by inhibition of CD28 and/or CTLA-4 interaction with CD80 (B7-1) and CD86 (B7-2). Disorders include immune system diseases (e.g. Autoimmune diseases such as psoriasis, multiple sclerosis, lupus erythematosus, diabetes, rheumatoid arthritis, and therapy against solid organ/cellular transplant rejection) and cancer.

XX

SQ Sequence 12 AA;

Query Match

Best Local Similarity 75.6%; Score 31; DB 20; Length 12;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 MYPTYLK 8

Db 2 mypyl 8

RESULT 11

AAV1222

ID AAV1222 standard; Protein; 58 AA.

XX

AC AAV1222;

XX

DT 18-JUN-1999 (first entry)

XX

DE Human 5' EST secreted protein SEQ ID NO: 535.

XX

Human 72 kDa gelatinase catalytic domain protein sequence SEQ ID NO:8.

Human; 92 kDa gelatinase; 72 kDa gelatinase; matrilysin; gelatinase;

catalytic domain; 92 CD; 72 CD; matrilysin CD; elastin; vulnary;

Human; secreted protein; EST; expressed sequence tag; diagnosis; forensic; gene therapy; chromosome mapping; signal peptide; upstream regulatory sequence; cytokine activity; cell proliferation; differentiation; haematopoiesis regulation; tissue growth regulation; reproductiv hormone regulation; chemotactic; chemokinetic; haemostatic; thrombolytic; anti-inflammatory; tumour inhibition.

OS Homo sapiens.

XX

PN WO9906554-A2.

XX

PD 11-FEB-1999.

XX

PF 31-JUL-1998; 98WO-IB01238.

XX

PR 01-AUG-1997; 97US-0905134.

XX

PA (GEST) GENSET.

XX

PI Duclert A, Dumas Milne Edwards J, Lacroix B;

XX

WPI; 1999-153784/13.

XX

N-PSDB; AAX41055.

XX

New nucleic acids encoding human secreted proteins - obtained from cDNA libraries prepared from kidney, fetal kidney, dystrophic muscle, muscle and heart tissue

XX

Claim 34; Page 598; 622pp; English.

XX

AAW40826 to AAX41093 represent 5' expressed sequence tags (ESTs) for human secreted proteins, and encode the proteins given in AAY01602 and AAY11994 to AAY12260, respectively. The proteins given represent the signal peptide and an N-terminal fragment of a secreted protein. The nucleic acid sequences can be used for producing secreted human gene products. They can also be used to develop products for diagnosis and therapy. The proteins obtained may have cytokine activity, cell proliferation/differentiation activity, haematopoiesis regulating activity, tissue growth regulating activity, reproductive hormone regulating activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, receptor/ligand activity, anti-inflammatory activity, tumour inhibition activity or other activities. The products can be used in forensic, gene therapy and chromosome mapping procedures. The sequences can also be used for obtaining corresponding promoter sequences. The nucleic acids encoding the signal peptide can be used for directing extracellular secretion of a polypeptide or the insertion of a polypeptide into a membrane, or importing a polypeptide into a cell.

XX

SQ Sequence 58 AA;

Query Match

Best Local Similarity 75.6%; Score 31; DB 20; Length 58;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYL 7

Db 1 mypsyl 6

RESULT 12

AAW74618

ID AAW74618 standard; protein; 173 AA.

XX

AC AAW74618;

XX

DT 22-MAY-2001 (first entry)

XX

Human 72 kDa gelatinase catalytic domain protein sequence SEQ ID NO:8.

XX

Human; 92 kDa gelatinase; 72 kDa gelatinase; matrilysin; gelatinase;

KW

catalytic domain; 92 CD; 72 CD; matrilysin CD; elastin; vulnary;

KW excess connective tissue removal; dermatological; keloid; scleroderma;
 KW post-operative fibrosis; intervertebral disc injection; fibrotic disease;
 KW hypertrophic scar; wound debridement; post-surgical adhesion;
 XX idiopathic pulmonary fibrosis.

OS Homo sapiens.

PN US6194189-B1.

XX 27-FEB-2001.

XX 16-DEC-1994; 94US-0357820.

XX 16-DEC-1994; 94US-0357820.

XX (UNIW) UNIV WASHINGTON.

XX Senior RM;

XX WPI; 2001-243407/25.

DR New gelatinase truncated mutant useful for treating disorders requiring
 PT the removal of excess connective tissue, e.g. keloids, post-operative
 PT fibrosis, intervertebral disc injections, hypertrophic scars -

XX Example; Column 13-16; l1pp; English.

CC The present invention describes a cDNA sequence which encodes the
 CC truncated mutant of the 92 kDa gelatinase having an amino acid sequence
 CC as given in AAB74617 consisting of residues 106-216 fused to residues
 CC 391-443 of the parent molecule. The truncated mutant has dermatological
 CC and vulnary activities. The truncated mutant is useful for treating
 CC disorders requiring the removal of excess connective tissue,
 CC e.g., keloids, post-operative fibrosis, intervertebral disc injections,
 CC hypertrophic scars, wound debridement, post-surgical adhesions and
 CC various fibrotic diseases (including scleroderma, idiopathic pulmonary
 CC fibrosis). The truncated mutant is catalytically active compared to
 CC the full protein. Unlike the full protein, the truncated mutant is
 CC essentially inactive against insoluble elastin, and does not require
 CC activation to be enzymatically active. The present sequence represents
 CC a 72 kDa gelatinase catalytic domain (72 CD) containing protein, which
 CC is given in the exemplification of the present invention.

XX Sequence 173 AA;

Query Match 75.6%; Score 31; DB 22; Length 173;
 Best Local Similarity 100.0%; Pred. No. 89;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6

|||||

DB 138 mypty 142

RESULT 13

ID AAB68682

AC AAB68682 standard; Protein; 173 AA.

XX AAB68682;

XX 02-MAY-2001 (first entry)

XX Matrilysin catalytic domain.

XX Gelatinase; excess connective tissue removal; matrilysin;

KW matrix metalloproteinase.

XX Unidentified.

XX US6184021-B1.

XX 06-FEB-2001.

XX 19-MAY-1995; 95US-0444628.

XX 16-DEC-1994; 94US-0357820.

XX (UNIW) UNIV WASHINGTON.

XX Senior RM;

XX WPI; 2001-202001/20.

XX New truncated mutant of 92 kDa gelatinase which is catalytically
 PT active, but is inactive against insoluble elastin, useful for treating
 PT disorders requiring the removal of excess connective tissues such as
 PT keloids -

XX Examples; Fig 2; l1pp; English.

XX The present invention relates to a truncated mutant (92 CD) of the 92
 CC kilo Dalton (kDa) gelatinase (see AAB68681). The truncated protein is
 CC useful for treating disorders requiring the removal of excess connective
 CC tissue, e.g. keloids, post-operative fibrosis, intervertebral disc
 CC injections, hypertrophic scars, wound debridement, post-surgical
 CC adhesions and various fibrotic diseases (scleroderma, idiopathic
 CC pulmonary fibrosis). Gelatinase is a matrix metalloproteinase and is also
 CC known as gelatinase B and MMP-9. The truncated protein is catalytically
 CC active comparable to the full protein but unlike the full protein is
 CC essentially inactive against insoluble elastin. The present sequence is
 CC the catalytic domain of matrilysin, which was used in a sequence homology
 CC alignment with the 92 CD protein of the present invention. Matrilysin is
 CC also a matrix metalloproteinase.

XX Sequence 173 AA;

Query Match 75.6%; Score 31; DB 22; Length 173;

Best Local Similarity 100.0%; Pred. No. 89;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6

|||||

DB 138 mypty 142

RESULT 14

AAW76253

ID AAW76253 standard; Protein; 261 AA.

XX AAW76253;

XX 02-DEC-1998 (first entry)

XX Human matrilysin-like protein.

XX Matrilysin; inhibitor; tumour necrosis factor; screening; MMP; infection;
 KW matrix metalloproteinase; antagonist; diagnosis; underexpression; cancer;
 KW degeneration; extracellular matrix; arthritis; cardiovascular disease;
 KW cachexia; multiple sclerosis.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..22

FT Protein /label= signal

FT Protein 23..261

FT /label= matrilysin_like_protein

XX WO9831818-A2.

XX 23-JUL-1998.

XX 20-JAN-1998; 98WO-US00783.

XX

PR 01-AUG-1997; 97US-0054541.
 PR 21-JAN-1997; 97US-0034205.
 PR 13-JUN-1997; 97US-0049607.

XX (HUMA-) HUMAN GENOME SCI INC.

PI Brewer L, Gentz R, Ni J, Rosen CA, Ruben SM;

XX WPI: 1998-414114/35.

DR N-PSDB; RAV61633.

XX Isolated nucleic acid encoding human metallo-protease(s) - used for
 PT diagnosis, treatment and prevention of, e.g. cancer, inflammation,
 PT neurological disease and infections

XX Claim 20a: Fig 4; 8lpp; English.

XX This sequence represents a novel human matrilysin-like protein which is
 CC an inhibitor of the members of the matrix metalloproteinase, MMP,
 CC family. This protein can be used in assays to screen for agonists and
 CC antagonists and the nucleic acid is used as a probe for gene mapping, in
 CC situ hybridisation and detection of corresponding genes in human tissue,
 CC and as sources of probes and primers for diagnosis. The protein and its
 CC antigenic fragments are used to raise antibodies (Ab) (which can be used
 CC for diagnosis in usual immunoassays or for in vivo imaging) and to screen
 CC for (ant)agonists. Agonists can be used to treat cancer or other
 CC conditions associated with underexpression of TNF-alpha. Antagonists are
 CC used to treat diseases associated with degeneration of the extracellular
 CC matrix (cancer, arthritis, cardiovascular disease, cachexia and multiple
 CC sclerosis).

XX Sequence 261 AA;

Query Match 75.6%; Score 31; DB 19; Length 261;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 MYPTY 6

Db 226 mypty 230

RESULT 15

AAW76254
 ID AAW76254 standard; Protein; 264 AA.

XX AAW76254;

02-DEC-1998 (first entry)

XX Human matrilysin protein.

XX Matrilysin; inhibitor; tumour necrosis factor; screening; MMP; infection;
 KW matrix metalloproteinase; antagonist; diagnosis; underexpression; cancer;
 KW degeneration; extracellular matrix; arthritis; cardiovascular disease;
 KW cachexia; multiple sclerosis.

XX Homo sapiens.

XX WO9831818-A2.

XX 23-JUL-1998.

XX 20-JAN-1998; 98WO-US00783.

XX 01-AUG-1997; 97US-0054541.

XX 21-JAN-1997; 97US-0034205.

XX 13-JUN-1997; 97US-0049607.

XX (HUMA-) HUMAN GENOME SCI INC.

PI Brewer L, Gentz R, Ni J, Rosen CA, Ruben SM;

XX WPI: 1998-414114/35.

XX Isolated nucleic acid encoding human metallo-protease(s) - used for
 PT diagnosis, treatment and prevention of, e.g. cancer, inflammation,
 PT neurological disease and infections

XX Disclosure: Fig 5; 8lpp; English.

XX This sequence represents the human matrilysin protein which is an
 CC inhibitor of the members of the matrix metalloproteinase, MMP, family.
 CC This protein is used in the identification of a novel human
 CC matrilysin-like protein which can be used in assays to screen for
 CC agonists and antagonists and the nucleic acid is used as a probe for gene
 CC mapping, in situ hybridisation and detection of corresponding genes in
 CC human tissue, and as sources of probes and primers for diagnosis. The
 CC protein and its antigenic fragments are used to raise antibodies (Ab)
 CC (which can be used for diagnosis in usual immunoassays or for in vivo
 CC imaging) and to screen for (ant)agonists. Agonists can be used to treat
 CC cancer or other conditions associated with underexpression of TNF-alpha.
 CC Antagonists are used to treat diseases associated with degeneration of
 CC the extracellular matrix (cancer, arthritis, cardiovascular disease,
 CC cachexia and multiple sclerosis).

XX Sequence 264 AA;

Query Match 75.6%; Score 31; DB 19; Length 264;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 MYPTY 6

Db 232 mypty 236

Search completed: September 4, 2001, 15:50:25
 Job time: 163 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2001, 15:51:12 ; Search time 40.98 seconds
(without alignments)
4.020 Million cell updates/sec

Title: US-09-630-345-2
Perfect score: 41
Sequence: 1 XMYPTYLK 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2.6/ptodata/2/iaa/5A-COMB.pep:*
2: /cgn2.6/ptodata/2/iaa/5B-COMB.pep:*
3: /cgn2.6/ptodata/2/iaa/6A-COMB.pep:*
4: /cgn2.6/ptodata/2/iaa/6B-COMB.pep:*
5: /cgn2.6/ptodata/2/iaa/PTUS-COMB.pep:*
6: /cgn2.6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	97.6	8	1	US-07-620-410-3
2	31	75.6	8	2	US-08-922-267A-65
3	31	75.6	173	4	US-08-444-628-8
4	31	75.6	173	4	US-08-357-820-8
5	31	75.6	261	3	US-09-009-156-5
6	31	75.6	264	3	US-09-009-156-6
7	31	75.6	267	4	US-08-448-489-18
8	31	75.6	271	3	US-08-896-062-2
9	31	75.6	459	1	US-08-630-592-2
10	31	75.6	459	1	US-08-714-991-2
11	31	75.6	459	3	US-09-032-365A-2
12	31	75.6	460	1	US-08-630-592-7
13	31	75.6	460	1	US-08-714-991-7
14	31	75.6	460	3	US-09-032-365A-8
15	31	75.6	462	4	US-08-068-392-3
16	31	75.6	462	4	US-08-396-988-3
17	31	75.6	505	1	US-08-631-200-2
18	31	75.6	505	1	US-08-630-592-4
19	31	75.6	505	1	US-08-714-991-4
20	31	75.6	505	2	US-08-829-553-2
21	31	75.6	505	2	US-08-922-267A-2
22	31	75.6	505	2	US-08-936-706A-2
23	31	75.6	505	2	US-08-936-706A-2
24	31	75.6	505	3	US-09-248-203-2
25	31	75.6	505	3	US-09-032-365A-4
26	31	75.6	505	4	US-08-812-824-3
27	31	75.6	505	4	US-09-406-071-2

28	31	75.6	506	1	US-08-631-200-8	Sequence 8, Appli
29	31	75.6	506	2	US-08-829-553-8	Sequence 8, Appli
30	31	75.6	506	2	US-08-922-267A-8	Sequence 8, Appli
31	31	75.6	506	2	US-08-936-706A-8	Sequence 8, Appli
32	31	75.6	506	2	US-08-936-706A-8	Sequence 8, Appli
33	31	75.6	506	3	US-09-248-203-8	Sequence 8, Appli
34	31	75.6	506	3	US-09-032-365A-62	Sequence 62, Appli
35	31	75.6	506	4	US-08-812-824-4	Sequence 4, Appli
36	31	75.6	506	4	US-09-406-071-8	Sequence 8, Appli
37	31	75.6	512	3	US-09-032-365A-60	Sequence 60, Appli
38	31	75.6	518	3	US-09-032-365A-58	Sequence 58, Appli
39	31	75.6	561	1	US-08-714-991-27	Sequence 27, Appli
40	31	75.6	561	3	US-09-032-365A-10	Sequence 10, Appli
41	30	73.2	73	1	US-08-379-538-7	Sequence 7, Appli
42	30	73.2	452	2	US-08-731-079A-2	Sequence 2, Appli
43	30	73.2	547	4	US-09-178-252-11	Sequence 11, Appli
44	30	73.2	725	1	US-08-448-170-4	Sequence 4, Appli
45	30	73.2	725	4	US-08-961-803-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1
US-07-620-410-3
; Sequence 3, Application US/07620410
; Patent No. 5217953
; GENERAL INFORMATION:
; APPLICANT: Gozes, Iilana
; APPLICANT: Brennenman, Douglas E.
; APPLICANT: Fridkin, Mati
; APPLICANT: Moody, Terry
; TITLE OF INVENTION: A NOVEL VASOACTIVE INTESTINAL PEPTIDE
; TITLE OF INVENTION: ANTAGONIST
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
; STREET: 1615 L. Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/620,410
; FILING DATE: 19901130
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Scott, Watson T.
; REGISTRATION NUMBER: 26,581
; REFERENCE/DOCKET NUMBER: WTS/5683/82679
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 248453 CUSH
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; US-07-620-410-3

Query Match 97.6%; Score 40; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
|||||||

Db 2 MYPTLK 8

RESULT 2
US-08-922-267A-65
; Sequence 65, Application US/08922267A
; Patent No. 5861239
; GENERAL INFORMATION:
; APPLICANT: Kleya, Patrick W.
; APPLICANT: Moore, Karen J.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; NUMBER OF SEQUENCES: 82
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/922.267A
; FILING DATE: 2-SEP-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/829,553
; FILING DATE: 28-MAR-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/631,200
; FILING DATE: 12-APR-1996
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-085
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-922-267A-65

Query Match 75.6%; Score 31; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 MYPTY 6
Db 2 MYPTY 6

RESULT 3
US-08-444-628-8
; Sequence 8, Application US/08444628
; Patent No. 6184021
; GENERAL INFORMATION:
; APPLICANT: Senior, Robert M.
; TITLE OF INVENTION: Catalytically-active Gelatinase Mutant
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Scott J. Meyer, Monsanto/Searle, A3SG
; STREET: 800 N. Lindbergh Blvd.
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63167
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/444.628
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/357,820
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyer, Scott J.
; REGISTRATION NUMBER: 25,275
; REFERENCE/DOCKET NUMBER: WU-2855
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)694-3117
; TELEFAX: (314)694-5435
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 173 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-444-628-8

Query Match 75.6%; Score 31; DB 4; Length 173;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 MYPTY 6
Db 138 MYPTY 142

RESULT 4
US-08-357-820-8
; Sequence 8, Application US/08357820
; Patent No. 6194189
; GENERAL INFORMATION:
; APPLICANT: Senior, Robert M.
; TITLE OF INVENTION: Catalytically-active Gelatinase Mutant
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scott J. Meyer, Monsanto/Searle, A3SG
; STREET: 800 N. Lindbergh Blvd.
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63167
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/357,820
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyer, Scott J.
; REGISTRATION NUMBER: 25,275
; REFERENCE/DOCKET NUMBER: WU-2855
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (314)694-3117
TELEFAX: (314)694-5435
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 173 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-357-820-8

Query Match 75.6%; Score 31; DB 4; Length 173;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
Db 138 MYPTY 142

ULT 5
US-09-009-156-5
Sequence 5, Application US/09009156
Patent No. 6046031
GENERAL INFORMATION:
APPLICANT: Ni, Jian
APPLICANT: Ruben, Steven
APPLICANT: Brewer, Laurie
APPLICANT: Gentz, Reiner
TITLE OF INVENTION: No. 6046031el Metalloproteinases
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: US
ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/009,156
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/034,205
FILING DATE: 21-JAN-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/049,607
FILING DATE: 13-JUN-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/054,541
FILING DATE: 01-AUG-1997
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PF376
TELEPHONE: 301-309-8504
TELEFAX: 301-309-8439
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 261 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-009-156-5

Query Match 75.6%; Score 31; DB 3; Length 261;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
Db 226 MYPTY 230

RESULT 6
US-09-009-156-6
Sequence 6, Application US/09009156
Patent No. 6046031
GENERAL INFORMATION:
APPLICANT: Ni, Jian
APPLICANT: Ruben, Steven
APPLICANT: Brewer, Laurie
APPLICANT: Gentz, Reiner
TITLE OF INVENTION: No. 6046031el Metalloproteinases
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: US
ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/009,156
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/034,205
FILING DATE: 21-JAN-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/049,607
FILING DATE: 13-JUN-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/054,541
FILING DATE: 01-AUG-1997
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PF376
TELEPHONE: 301-309-8504
TELEFAX: 301-309-8439
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 264 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-009-156-6

Query Match 75.6%; Score 31; DB 3; Length 264;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
Db 232 MYPTY 236

RESULT 7

US-08-448-489-18
; Sequence 18, Application US/08448489
; Patent No. 6184022
; GENERAL INFORMATION:
; APPLICANT: SEIKI Motoharu
; APPLICANT: SATO, Hiroshi
; APPLICANT: SHINAGAWA, Akira
; TITLE OF INVENTION: NOVEL METALLOPROTEINASE AND ENCODING DNA THEREFOR
; FILE REFERENCE: 55-290P
; CURRENT APPLICATION NUMBER: US/08/448,489
; CURRENT FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Known Member of
; OTHER INFORMATION: Matrix Metalloproteinase Family
08-448-489-18

Query Match 75.6%; Score 31; DB 4; Length 267;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
Db 232 MYPTY 236

RESULT 8
US-08-896-062-2
; Sequence 2, Application US/08896062
; Patent No. 6010893
; GENERAL INFORMATION:
; APPLICANT: KIHARA, Yasunori
; TITLE OF INVENTION: PROCESS FOR PRODUCING HUMAN MATRILYSIN
; TITLE OF INVENTION: BY MEANS OF RECOMBINANT DNA
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: 419 Broadway and Nelmark
; STREET: 419 Seventh Street N.W., Ste. 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States of America
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/896,062
; FILING DATE: 17-JUL-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/530,984
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Browdy, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: KIHARA-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-5197
; TELEFAX: (202) 737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 271 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

; MOLECULE TYPE: protein
US-08-896-062-2

Query Match 75.6%; Score 31; DB 3; Length 271;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
Db 236 MYPTY 240

RESULT 9
US-08-630-592-2
; Sequence 2, Application US/08630592
; Patent No. 5770432
; GENERAL INFORMATION:
; APPLICANT: Nishina, Patsy
; APPLICANT: NO. 5770432entrauth, Konrad
; APPLICANT: Nagger, Juergeen
; APPLICANT: NO. 5770432th, Michael
; TITLE OF INVENTION: Obesity Associated Genes
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
; STREET: 3400 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MSDOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08/630,592
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J.
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: A59504/BIR/PJS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 7811989
; TELEFAX: (415) 3983249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 459 amino acids
; TYPE: amino acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-630-592-2

Query Match 75.6%; Score 31; DB 1; Length 459;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
Db 233 MYPTY 237

RESULT 10
US-08-714-991-2
; Sequence 2, Application US/08714991
; Patent No. 5776762
; GENERAL INFORMATION:
; APPLICANT: NORTH, Michael

APPLICANT: NISHINA, Patsy
APPLICANT: NO. 5776762en-Trauth, Konrad
APPLICANT: NAGGERT, Juergen
TITLE OF INVENTION: OBESITY ASSOCIATED GENES
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/714,991
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: SHERWOOD, Pamela J.
REGISTRATION NUMBER: 36,677
REFERENCE/DOCKET NUMBER: A-59504-1/PJS
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-494-8700
TELEFAX: 415-494-8771
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 459 amino acids
TYPE: amino acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-714-991-2

Query Match 75.6%; Score 31; DB 1; Length 459;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
|||||
Db 233 MYPTY 237

RESULT 11
US-09-032-365A-2
Sequence 2, Application US/09032365A
Patent No. 6114502
GENERAL INFORMATION:
APPLICANT: No. 6114502th, Michael
APPLICANT: Nishina, Patsy
APPLICANT: NAGGERT, Juergen
APPLICANT: NO. 6114502en-Trauth, Konrad
TITLE OF INVENTION: GENE FAMILY ASSOCIATED WITH
TITLE OF INVENTION: NEUROSENSORY DEFECTS
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bozicevic & Reed, LLP
STREET: 285 Hamilton Avenue, Suite 200
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/032,365A
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Sherwood, Pamela J.
REGISTRATION NUMBER: 36,677
REFERENCE/DOCKET NUMBER: SEQ-2CIP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-327-3400
TELEFAX: 650 327-3231
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 459 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-032-365A-2

Query Match 75.6%; Score 31; DB 3; Length 459;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
|||||
Db 233 MYPTY 237

RESULT 12
US-08-630-592-7
Sequence 7, Application US/08630592
Patent No. 5770432
GENERAL INFORMATION:
APPLICANT: Nishina, Patsy
APPLICANT: NO. 5770432enTrauth, Konrad
APPLICANT: Naggert, Juergen
APPLICANT: NO. 5770432th, Michael
TITLE OF INVENTION: Obesity Associated Genes
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
STREET: 3400 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 941114187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/630,592
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sherwood, Pamela J.
REGISTRATION NUMBER: 36,677
REFERENCE/DOCKET NUMBER: A59504/BIR/PJS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 7811989
TELEFAX: (415) 3983249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 460 amino acids
TYPE: amino acid
STRANDEDNESS: double

TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-630-592-7

Query Match 75.6%; Score 31; DB 1; Length 460;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
Db 234 MYPTY 238

RESULT 13
US-08-714-991-7
Sequence 7, Application US/08714991
Patent No. 5776762
GENERAL INFORMATION:
APPLICANT: NORTH, Michael
APPLICANT: NISHINA, Patsy
APPLICANT: NAGERT, Juergen
TITLE OF INVENTION: OBESITY ASSOCIATED GENES
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HUBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarradero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/714,991
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: SHERWOOD, Pamela J.
REGISTRATION NUMBER: 36,677
REFERENCE/DOCKET NUMBER: A-59504-1/PJS
TELEPHONE: 415-494-8771
TELEFAX: 415-494-8771
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 460 amino acids
TYPE: amino acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-714-991-7

Query Match 75.6%; Score 31; DB 1; Length 460;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
Db 234 MYPTY 238

RESULT 14
US-09-032-365A-8
Sequence 8, Application US/09032365A
Patent No. 6114502
GENERAL INFORMATION:

APPLICANT: No. 6114502th, Michael
APPLICANT: Nishina, Patsy
APPLICANT: Nagert, Juergen
APPLICANT: No. 6114502en-Trauth, Konrad
TITLE OF INVENTION: GENE FAMILY ASSOCIATED WITH
NEUROSENSORY DEFECTS
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bozicevic & Reed, LLP
STREET: 285 Hamilton Avenue, Suite 200
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/032,365A
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: SHERWOOD, Pamela J.
REGISTRATION NUMBER: 36,677
REFERENCE/DOCKET NUMBER: SEQ-2CIP2
TELEPHONE: 650-327-3400
TELEFAX: 650-327-3231
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 460 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-032-365A-8

Query Match 75.6%; Score 31; DB 3; Length 460;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
Db 234 MYPTY 238

RESULT 15
US-08-068-392-3
Sequence 3, Application US/08068392
Patent No. 6150152
GENERAL INFORMATION:
APPLICANT: Shapiro, Steven M.
TITLE OF INVENTION: Human Macrophage Metalloproteinase
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scott J. Meyer, Monsanto Co., A3SM
STREET: 800 N. Lindbergh Blvd.
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63167
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/068,392
;; FILING DATE: 19930528
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Meyer, Scott J.
;; REGISTRATION NUMBER: 25275
;; REFERENCE/DOCKET NUMBER: 07-24(12406)A
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (314)694-3117
;; TELEFAX: (314)694-5435
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 462 amino acids
;; TYPE: AMINO ACID
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-08-068-392-3

Query Match 75.6%; Score 31; DB 4; Length 462;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
|
Db 229 MYPTY 233

Search completed: September 4, 2001, 15:51:12
Job time: 195 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2001, 15:52:08 ; Search time 49.08 Seconds
(without alignments)
12.416 Million cell updates/sec

Title: US-09-630-345-2

Perfect score: 41

Sequence: 1 MYPTYLK 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_68:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match %	Length	DB ID	Description
1	37	90.2	384	2 IS1267	collagenase (EC 3.4.24.-) - bullfrog
2	35	85.4	393	2 B86644	transporter ybfb [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
3	34	82.9	436	2 C69764	4-aminobutyrate am
4	33	80.5	218	2 E86750	hypothetical prote
5	33	80.5	269	2 D69011	hypothetical prote
6	32	78.0	340	1 MMBE5	cell fusion protei
7	32	78.0	349	2 JU0458	peroxidase (EC 1.1
8	32	78.0	593	2 JC4884	organic cation tra
9	32	78.0	1263	2 F44466	DNA-directed RNA p
10	31	75.6	247	2 B85515	hypothetical prote
11	31	75.6	262	2 T29698	hypothetical prote
12	31	75.6	267	1 KCHUM	matrilysin (EC 3.4
13	31	75.6	267	2 A57490	matrilysin (EC 3.4
14	31	75.6	301	2 S57923	SEC14 protein - ye
15	31	75.6	378	2 F86307	hypothetical prote
16	31	75.6	385	2 E86359	Similar to seed ma
17	31	75.6	462	2 A42401	macrophage elastas
18	31	75.6	478	2 A86677	phospho-beta-gluc
19	31	75.6	483	2 JC5743	matrix metalloprot
20	31	75.6	495	1 PLWLB	L1 protein - bovin
21	31	75.6	501	1 PLWLB2	L1 protein - bovin
22	31	75.6	501	1 PLWLEP	L1 protein - Europ
23	31	75.6	502	2 T19708	hypothetical prote
24	31	75.6	503	2 T23175	hypothetical prote
25	31	75.6	505	2 S68518	tub protein, brain
26	31	75.6	511	2 T19496	hypothetical prote
27	31	75.6	513	1 PLWLDP	L1 protein - deer
28	31	75.6	516	2 T15633	hypothetical prote
29	31	75.6	565	2 H69363	hypothetical prote

30	31	75.6	662	2 S42826	probable ATPase -
31	31	75.6	2185	1 GNNYSV	genome polyprotein
32	31	75.6	2185	1 GNNYSH	genome polyprotein
33	30	73.2	141	2 T29507	hypothetical prote
34	30	73.2	152	2 G83476	hypothetical prote
35	30	73.2	155	2 T17986	hypothetical prote
36	30	73.2	192	2 T11930	NADH dehydrogenase
37	30	73.2	239	2 B64757	glycolate oxidase
38	30	73.2	239	2 D85522	probable dehydroge
39	30	73.2	248	2 S57910	probable response
40	30	73.2	262	2 C81384	shikimate 5-dehydr
41	30	73.2	287	2 D83414	probable transcrip
42	30	73.2	343	1 MMBEA5	cell fusion protei
43	30	73.2	368	1 HLBECH	membrane glycoprot
44	30	73.2	410	2 T20397	hypothetical prote
45	30	73.2	491	2 JE0396	phospho-beta-galac

ALIGNMENTS

RESULT 1
IS1267
collagenase (EC 3.4.24.-) - bullfrog
C:Species: Rana catesbeiana (bullfrog)
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 22-Jun-1999
C:Accession: IS1267
R:Oofusa, K.; Yomori, S.; Yoshizato, K.
Int. J. Dev. Biol. 38, 345-350, 1994
A:Title: Regionally and hormonally regulated expression of genes of collagen and coll
A:Reference number: IS1267; MUID:95071832
A:Accession: IS1267
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-384 <OOF>
A:Cross-references: GB:S75623; NID:g913070; PIDN:AAB32661.1; PID:g913071
C:Superfamily: interstitial collagenase; hemopexin repeat homology; matrix metallopro
F:Keywords: hydrolase; metalloproteinase; zinc; zymogen
F:53-231/Domain: matrix metalloproteinase homology <MMP>
F:236-381/Domain: hemopexin repeat homology <PXN>
F:81,189,193,199/Binding site: zinc, catalytic (Cys, His, His, His) (inhibited) #stat
F:189,193,199/Binding site: zinc, catalytic (His) (active) #status predicted
F:190/Active site: Glu #status predicted

Query Match 90.2%; Score 37; DB 2; Length 384;
Best Local Similarity 85.7%; Pred. No. 5;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 2 MYPTYLK 8
DB 207 MYPTYLK 213
|||||:
B86644
RESULT 2
B86644
transporter ybfb [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 23-Mar-2001
C:Accession: B86644
R:Boletín, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissensbach, J.; Eh
Genome Res. in press, 2001
A:Title: The complete genome sequence of the lactic acid bacterium.
A:Reference number: A86625
A:Accession: B86644
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-393 <STO>
A:Cross-references: GB:AE005176; NID:gl2723004; PIDN:AAK04252.1; GSPDB:GN00146
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: ybfb

Query Match 85.4%; Score 35; DB 2; Length 393;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYL 7
 |||||
 Db 278 MYPTYL 283

RESULT 3
 C69764
 4-aminobutyrate aminotransferase homolog ycnG - Bacillus subtilis
 C:Species: Bacillus subtilis
 C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
 R:Accession: C69764
 R:Authors: Ogawara, N.; Moszer, I.; Albertini, A.M.; Allioni, G.; Azevedo, V.; Berteu
 C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
 C.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
 are 390, 249-256, 1997
 Authors: Foulger, D.; Fritz, C.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
 lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
 Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois
 A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauee
 Y. M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Satoh, T.; Scanlon
 A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Serol
 akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Toononi, A.; Tosato, V.; Uchiyama
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
 A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Zumbstein, A.
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
 A:Reference number: A69580; MUID:98044033
 A:Accession: C69764
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-436 <NUN>
 A:Cross-references: GB:Z99106; GB:AL009126; NID:g2632653; PIDN:CAB12198.1; PID:g2632691
 A:Experimental source: strain 168
 C:Genetics:
 A:Gene: ycnG
 C:Superfamily: ornithine--oxo-acid aminotransferase

Query Match 82.9%; Score 34; DB 2; Length 436;
 Best Local Similarity 71.4%; Pred. No. 23;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYL 8
 |||||
 Db 88 MYPTYE 94

RESULT 4
 E86750
 hypothetical protein ykdb [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
 C:Species: Lactococcus lactis subsp. lactis
 C>Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 23-Mar-2001
 R:Accession: E86750
 R:Authors: Bolotin, A.; Wincker, P.; Mauger, S.; Jallion, O.; Malarne, K.; Weissensbach, J.; Ehrli
 Genome Res. in press, 2001
 A:Title: The complete genome sequence of the lactic acid bacterium.
 A:Reference number: A66625
 A:Accession: E86750
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-218 <STO>
 A:Cross-references: GB:AE005176; NID:g12723952; PIDN:AAK05103.1; GSPDB:GN00146
 A:Experimental source: strain IL1403
 C:Genetics:
 A:Gene: ykdb

Query Match 80.5%; Score 33; DB 2; Length 218;

Best Local Similarity 83.3%; Pred. No. 19;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 YPTYLK 8
 |||||
 Db 204 YPTYIK 209

RESULT 5
 D69011
 hypothetical protein MTH1087 - Methanobacterium thermoautotrophicum (strain Delta H)
 C:Species: Methanobacterium thermoautotrophicum
 C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
 R:Accession: D69011
 R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T
 ; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani,
 ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
 J. Bacteriol. 179, 7135-7155, 1997
 A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: fu
 A:Reference number: A69000; MUID:98037514
 A:Accession: D69011
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-269 <MTH>
 A:Cross-references: GB:AE000879; GB:AE000666; NID:g2622175; PIDN:AA85576.1; PID:g262
 A:Experimental source: strain Delta H
 C:Genetics:
 A:Gene: MTH1087

Query Match 80.5%; Score 33; DB 2; Length 269;
 Best Local Similarity 83.3%; Pred. No. 23;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 YPTYLK 8
 |||||
 Db 175 YPTYIK 180

RESULT 6
 MMBE5
 cell fusion protein precursor - human herpesvirus 3
 C:Species: human herpesvirus 3, varicella-zoster virus
 C>Date: 30-Sep-1988 #sequence_revision 30-Sep-1988 #text_change 16-Jul-1999
 R:Accession: E27212
 R:Davidson, A.J.; Scott, J.E.
 J. Gen. Virol. 67, 1759-1816, 1986
 A:Title: The complete DNA sequence of varicella-zoster virus.
 A:Reference number: A27345; MUID:86306657
 A:Accession: E27212
 A:Molecule type: DNA
 A:Residues: 1-340 <DAV>
 A:Cross-references: EMBL:X04370; NID:g59589; PIDN:CAA27888.1; PID:g59994
 C:Genetics:
 A:Gene: 5
 C:Superfamily: herpesvirus cell fusion protein
 C:Keywords: membrane fusion; transmembrane protein
 F; 1-31/Domain: signal sequence #status predicted <SIG>
 F; 32-340/Product: cell fusion protein #status predicted <CPF>
 F; 115-137/Domain: transmembrane #status predicted <TM1>
 F; 220-238/Domain: transmembrane #status predicted <TM3>
 F; 251-269/Domain: transmembrane #status predicted <TM4>
 F; 307-322/Domain: transmembrane #status predicted <TM5>

Query Match 78.0%; Score 32; DB 1; Length 340;
 Best Local Similarity 57.1%; Pred. No. 47;
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
 |||||
 Db 244 LYPTYIR 250

RESULT 7

JU0458
 peroxidase (EC 1.11.1.7) E - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 04-Mar-2000
 C:Accession: JU0458
 R:Itapiruk, C.; Higashimura, N.; Yamamoto, K.; Okada, N.; Shinmyo, A.; Takano, M.
 Gene 98, 237-241, 1991
 A:Title: Nucleotide sequences of two genomic DNAs encoding peroxidase of Arabidopsis thaliana

A:Reference number: JU0457; MUID:91200671
 A:Accession: JU0458
 A:Molecule type: DNA
 A:Residues: 1-349 <INT>
 A:Cross-references: GB:M58381; NID:g166806; PIDN:AAA32842.1; PID:g166807
 C:Genetics:
 A:Gene: prxEa
 A:Introns: 76/3; 140/3; 197/1
 C:Superfamily: peroxidase
 C:Keywords: chromoprotein; heme; iron; metalloprotein; oxidoreductase
 F:77/Active site: Arg #status predicted
 F:71,199/Binding site: heme iron (His) (axial ligands) #status predicted
 F:73-78/Disulfide bonds: #status predicted
 F:126-329/Disulfide bonds: #status predicted
 F:206-238/Disulfide bonds: #status predicted

Query Match 78.0%; Score 32; DB 2; Length 349;
 Best Local Similarity 83.3%; Pred. No. 48;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 MYPTYL 7

:|||||

Db 226 LYPTYL 231

RESULT 8

JC4884
 organic cation transporter protein 2 - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 10-Sep-1996 #sequence_revision 18-Oct-1996 #text_change 05-Nov-1999
 C:Accession: JC4884
 R:Okuda, M.; Saito, H.; Urakami, Y.; Takano, M.; Inui, K.
 Biochem. Biophys. Res. Commun. 224, 500-507, 1996
 A:Title: cDNA cloning and functional expression of a novel rat kidney organic cation transporter
 A:Reference number: JC4884; MUID:96295517
 A:Accession: JC4884

Molecule type: mRNA

A:Residues: 1-593 <OKU>
 A:Cross-references: DDBJ:DB3044; NID:g1502282; PIDN:BAAL1754.1; PID:d1012421; PID:g15022
 A:Experimental source: kidney
 C:Comment: This protein is responsible for the transport of cationic drugs in kidney.

Query Match 78.0%; Score 32; DB 2; Length 593;
 Best Local Similarity 57.1%; Pred. No. 82;
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 MYPTYLK 8

:|||||

Db 457 LYPTIIR 463

RESULT 9

F44466
 DNA-directed RNA polymerase (EC 2.7.7.6) beta chain - Thermotoga maritima (strain MSB8)
 C:Species: Thermotoga maritima
 C:Date: 10-Jun-1993 #sequence_revision 26-Apr-1996 #text_change 21-Jul-2000
 C:Accession: S41466; S41462; F44466; B72373; S19903
 R:Palmer, P.; Schleper, C.; Arnold-Ammer, I.; Holz, I.; Meier, T.; Lottspeich, F.; Zillig, G.
 Nucleic Acids Res. 21, 4904-4908, 1993
 A:Title: The DNA-dependent RNA-polymerase of Thermotoga maritima; characterisation of the

A:Reference number: S41462; MUID:94232816
 A:Accession: S41466

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-1263 <PAL>

A:Cross-references: EMBL:X72695; NID:g425255; PIDN:CAA51246.1; PID:g425258

A:Experimental source: DSM 3109

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1993

A:Accession: S41462

A:Molecule type: protein

A:Residues: 1-5, 'X', 7-23 <PAW>

A:Experimental source: DSM 3109

R:Liao, D.; Dennis, P.P.

J. Biol. Chem. 267, 22787-22797, 1992

A:Title: The organization and expression of essential transcription translation compo

A:Reference number: A44466; MUID:93054590

A:Accession: F44466

A:Molecule type: DNA

A:Residues: 1-404 <LIA>

A:Cross-references: EMBL:Z11839; NID:g407020; PIDN:CAA77863.1; PID:g48188

A:Note: sequence extracted from NCBI backbone (NCBIP:118059)

R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hic
 Garret, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson,
 C.M.

Nature 399, 323-329, 1999

A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome

A:Reference number: A72200; MUID:99287316

A:Accession: B72373

A:Molecule type: DNA

A:Residues: 1-1263 <ARN>

A:Cross-references: GB:AE001724; GB:AE000512; NID:g4980966; PIDN:AAD35543.1; PID:g49498

A:Experimental source: strain MSB8

C:Genetics:

A:Gene: rpoB

C:Superfamily: DNA-directed RNA polymerase beta chain

C:Keywords: nucleotidyltransferase; transcription

F:1-1263/Product: DNA-directed RNA polymerase beta chain #status predicted <MAT>

Query Match 78.0%; Score 32; DB 2; Length 1263;
 Best Local Similarity 83.3%; Pred. No. 1.8e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 MYPTYL 7

:|||||

Db 214 LYPTYL 219

RESULT 10

B85515

hypothetical protein 20321 [imported] - Escherichia coli (strain O157:H7)

C:Species: Escherichia coli

C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 31-Mar-2001

C:Accession: B85515

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May

Iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apoda

Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: B85515

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-247 <STO>

A:Cross-references: GB:AE005174; NID:g12513034; PIDN:AAG54582.1; GSPDB:GN00145; UWGP:

A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:

C:Superfamily: hypothetical protein f242a

Query Match 75.6%; Score 31; DB 2; Length 247;

Best Local Similarity 83.3%; Pred. No. 54;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 YPTYLK 8
 DB 76 YPSYLK 81

RESULT 11
 T29698
 hypothetical protein F31A3.2 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
 C:Accession: T29698
 R:Murray, J.; Le, T.T.
 A:Submitted to the EMBL Data Library, May 1996
 A:Description: The sequence of *C. elegans* cosmid F31A3.
 A:Reference number: Z20667
 A:Accession: T29698
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-262 <MUL>
 A:Cross-references: EMBL:U58742; PIDN:AAB36855.1; GSPDB:GN00028; CESP:F31A3.2
 A:Experimental source: strain Bristol N2; clone F31A3
 C:Genetics:
 A:Gene: CESP:F31A3.2
 A:Map position: X
 A:Introns: 23/3; 80/2; 155/2

Query Match 75.6%; Score 31; DB 2; Length 262;
 Best Local Similarity 71.4%; Pred. No. 58;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 MPTYLK 8
 DB 188 MYPVLE 194

RESULT 12
 KCHDM
 matrilysin (EC 3.4.24.23) precursor - human
 N:Alternate names: matrin; matrix metalloproteinase 7 (MMP7); probable metalloproteinase
 A:Contains: promatrilysin
 C:Species: Homo sapiens (man)
 C>Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 18-Jun-1999
 C:Accession: B28816; A60539; S24324
 R:Muller, D.; Quantin, B.; Gesnel, M.C.; Millon-Collard, R.; Abecassis, J.; Breathnach, Biochem. J. 253, 187-192, 1988
 A:Title: The collagenase gene family in humans consists of at least four members.
 A:Reference number: A90339; MUID:86339885
 A:Accession: B28816
 A:Molecule type: mRNA
 A:Residues: 1-267 <MUL>
 A:Cross-references: EMBL:X07819; NID:g35798; PIDN:CAA30678.1; PID:g35799
 R:Miyazaki, K.; Hattori, Y.; Umenishi, F.; Yasumitsu, H.; Umeda, M. Cancer Res. 50, 7758-7764, 1990
 A:Title: Purification and characterization of extracellular matrix-degrading metalloproteinase
 A:Reference number: A60539; MUID:91070531
 A:Accession: A60539
 A:Molecule type: protein
 A:Residues: 18-35, X1, 37-42 <MY>
 R:Marti, H.P.; McNeill, L.; Thomas, G.; Davies, M.; Lovett, D.H. Biochem. J. 285, 899-905, 1992
 A:Title: Molecular characterization of a low-molecular-mass matrix metalloproteinase secreted by human placental cells
 A:Reference number: S24324; MUID:92359961
 A:Accession: S24324
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-267 <MAR>
 A:Cross-references: EMBL:Z11887; NID:g35802; PIDN:CAA77942.1; PID:g35803
 A:Comment: This enzyme is similar in its activity to stromelysin and degrades various extracellular matrix components of types II, IV, IX, X, and XI.
 C:Comment: Matrilysin hydrolyzes peptide bonds in plasminogen to yield a fragment with a C-terminal lysine.
 C:Genetics:

A:Gene: GDB:MMP7; MPSL1
 A:Cross-references: GDB:125751; OMIM:178990
 A:Map position: 11q21-11q22
 C:Superfamily: matrilysin; matrix metalloproteinase homology
 C:Keywords: calcium; extracellular matrix; fibroblast; hydrolase; metalloproteinase;
 F:1-17/Domain: signal sequence #status predicted <SIG>
 F:18-267/Product: promatrilysin #status predicted <PRO>
 F:18-94/Domain: activation peptide #status predicted <ACT>
 F:55-259/Domain: matrix metalloproteinase homology <MMP>
 F:85-92/Region: autoinhibitory
 F:95-267/Product: matrilysin #status predicted <MAT>
 F:87, 214, 218, 224/Binding site: zinc, catalytic (Cys, His, His, His) (inhibited) #stat
 F:214, 218, 224/Binding site: zinc, catalytic (His) (active) #status predicted
 F:215/Active site: Glu #status predicted

Query Match 75.6%; Score 31; DB 1; Length 267;
 Best Local Similarity 100.0%; Pred. No. 59;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
 DB 232 MYPTY 236

RESULT 13
 A57490
 matrilysin (EC 3.4.24.23) precursor - rat
 N:Alternate names: matrix metalloproteinase 7 (MMP7)
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 08-Dec-1995 #sequence_revision 08-Dec-1995 #text_change 22-Jun-1999
 C:Accession: A57490
 R:Abramson, S.R.; Conner, G.E.; Nagase, H.; Neuhaus, I.; Woessner Jr., J.F. J. Biol. Chem. 270, 16016-16022, 1995
 A:Title: Characterization of rat uterine matrilysin and its cDNA. Relationship to human matrilysin
 A:Reference number: A57490; MUID:95332299
 A:Accession: A57490
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-267 <ABR>
 A:Cross-references: GB:L24374; NID:g402492; PIDN:AAA99432.1; PID:g402493
 C:Superfamily: matrilysin; matrix metalloproteinase homology
 C:Keywords: hydrolase; metalloproteinase; zinc; zymogen
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-267/Product: matrilysin #status predicted <MAT>
 F:58-262/Domain: matrix metalloproteinase homology <MMP>
 F:90, 217, 221, 227/Binding site: zinc, catalytic (Cys, His, His, His) (inhibited) #stat
 F:217, 221, 227/Binding site: zinc, catalytic (His) (active) #status predicted
 F:218/Active site: Glu #status predicted

Query Match 75.6%; Score 31; DB 2; Length 267;
 Best Local Similarity 100.0%; Pred. No. 59;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
 DB 235 MYPTY 239

RESULT 14
 S57923
 SEC14 protein - yeast (*Candida albicans*)
 C:Species: *Candida albicans*
 C>Date: 13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 26-Feb-1998
 C:Accession: S57923; S72193
 R:Monteoliva, L.; Sanchez, M.; Pla, J.; Gil, C.; Nombela, C. Submitted to the EMBL Data Library, September 1994
 A:Description: Characterisation of the *Candida albicans* SEC14 homolog gene.
 A:Reference number: S57923
 A:Accession: S57923
 A>Status: preliminary
 A:Molecule type: DNA

A;Residues: 1-301 <MON>
A;Cross-references: EMBL:X81937
R;Montecoliva, L.; Sanchez, M.; Pla, J.; Gil, C.; Nombela, C.
Yeast 12, 1097-1105, 1996
A;Title: Cloning of Candida albicans SEC14 gene homologue coding for a putative essential protein
A;Reference number: S72193; MUID:97051600
A;Accession: S72193
A;Molecule type: DNA
A;Residues: 1-301 <MON>
A;Cross-references: EMBL:X81937
A;Note: the authors translated the codon CTG for residue 180 as Ser
C;Genetics:
A;Gene: SEC14
C;Superfamily: cellular retinaldehyde-binding protein; cellular retinaldehyde-binding protein
F;59-266/Domain: cellular retinaldehyde-binding protein homology <CRB>

Query Match 75.6%; Score 31; DB 2; Length 301;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
2 MYPTY 6
|||||
Db 106 MYPTY 110

RESULT 15
F86307
hypothetical protein AAD50017.1 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence-revision 02-Mar-2001 #text-change 31-Mar-2001
C;Accession: F86307
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 815-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719
A;Accession: F86307
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-378 <STO>
A;Cross-references: GB:AE005172; NID:g5734752; PIDN:AAD50017.1; GSPDB:GN00141
C;Genetics:
Gap position: 1

Query Match 75.6%; Score 31; DB 2; Length 378;
Best Local Similarity 100.0%; Pred. No. 84;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 MYPTY 6
|||||
Db 118 MYPTY 122

Search completed: September 4, 2001, 15:52:08
Job time: 231 sec

THIS PAGE BLANK (USPTO)


```

FT METAL 189 189 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 190 190 BY SIMILARITY.
FT METAL 193 193 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 199 199 ZINC (CATALYTIC) (BY SIMILARITY).
FT DISULFID 242 381 PROBABLE.
SQ SEQUENCE 384 AA; 43582 MW; A5B5E2FB332239DF CRC64;

Query Match 90.2%; Score 37; DB 1; Length 384;
Best Local Similarity 85.7%; Pred. No. 2.9;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
DB 207 MYPTYL 213

RESULT 2
GABT_BACSU STANDARD; PRT; 436 AA.
P94427;
15-DEC-1998 (Rel. 37, Created)
15-DEC-1998 (Rel. 37, Last sequence update)
30-MAY-2000 (Rel. 39, Last annotation update)
DE PROBABLE 4-AMINO BUTYRATE AMINOTRANSFERASE (EC 2.6.1.19) (GAMMA-AMINO-
DE N-BUTYRATE TRANSAMINASE) (GABA TRANSAMINASE) (GLUTAMATE-SUCCINIC
DE SEMIALDEHYDE TRANSAMINASE) (GABA AMINOTRANSFERASE) (GABA-AT).
GN GABT.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=97124189; PubMed=8969502;
RA Yamane K., Kumano M., Kurita K.;
RT "The 25 degrees-36 degrees region of the Bacillus subtilis
RT chromosome: determination of the sequence of a 146 kb segment and
RT identification of 113 genes.";
RL Microbiology 142:3047-3056(1996).
CC -!- CATALYTIC ACTIVITY: 4-AMINO BUTANOATE + 2-OXOGLUTARATE -> SUCCINATE
CC -!- SEMIALDEHYDE + L-GLUTAMATE.
CC -!- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -!- PATHWAY: 4-AMINO BUTYRATE (GABA) DEGRADATION PATHWAY.
CC -!- SIMILARITY: BELONGS TO CLASS-III OF PYRIDOXAL-PHOSPHATE-DEPENDENT
CC AMINOTRANSFERASES.

-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; D50453; BAA09021.1; -
DR EMBL; Z99108; CAB2198.1; -
DR Subtilisin; BGI2043; gabt.
DR InterPro; IPR000954; -
DR Pfam; PF00202; aminotran_3; 1.
DR PROSITE; PS00600; AA-TRANSFER_CLASS_3; 1.
KW Transferase; Aminotransferase; Pyridoxal phosphate.
FT BINDING 281 281 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ SEQUENCE 436 AA; 47249 MW; D0961F6D4189A8F3 CRC64;

Query Match 82.9%; Score 34; DB 1; Length 436;
Best Local Similarity 71.4%; Pred. No. 13;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
DB 207 MYPTYL 213

```

```

Db 88 MYPTYL 94

RESULT 3
CELL_VZVD STANDARD; PRT; 340 AA.
AC P09261;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE CELL FUSION PROTEIN PRECURSOR.
GN 5
OS Varicella-zoster virus (strain Dumas) (VZV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10338;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86306657; PubMed=3018124;
RA Davison A.J., Scott J.E.;
RT "The complete DNA sequence of varicella-zoster virus.";
RL J. Gen. Virol. 67:1759-1816(1986).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X04370; CAA27888.1; -
DR PIR; E27212; MMBES.
DR InterPro; IPR002587; -
DR Pfam; PF01621; Fusion_gly_K; 1.
KW Fusion protein; Transmembrane; Signal.
FT SIGNAL 1 ?
FT CHAIN ? 340 CELL FUSION PROTEIN.
FT CARBOHYD 55 55 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 65 65 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 340 AA; 38576 MW; 0387FE0EC39C946 CRC64;

Query Match 78.0%; Score 32; DB 1; Length 340;
Best Local Similarity 57.1%; Pred. No. 25;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
DB 244 LYPTYL 250

-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; D50453; BAA09021.1; -
DR EMBL; Z99108; CAB2198.1; -
DR Subtilisin; BGI2043; gabt.
DR InterPro; IPR000954; -
DR Pfam; PF00202; aminotran_3; 1.
DR PROSITE; PS00600; AA-TRANSFER_CLASS_3; 1.
KW Transferase; Aminotransferase; Pyridoxal phosphate.
FT BINDING 281 281 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ SEQUENCE 436 AA; 47249 MW; D0961F6D4189A8F3 CRC64;

Query Match 82.9%; Score 34; DB 1; Length 436;
Best Local Similarity 71.4%; Pred. No. 13;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
DB 207 MYPTYL 213

```

Gene 98:237-241(1991).

RL -1- FUNCTION: REMOVAL OF H₂O₂(2), OXIDATION OF TOXIC REDUCTANTS,
 CC BIOSYNTHESIS AND DEGRADATION OF LIGNIN, DEFENSE RESPONSE TOWARD
 CC WOUNDING ON METABOLISM OF AUXIN. THESE FUNCTIONS MIGHT BE
 CC DEPENDENT ON EACH ISOZYME/ISOFORM IN EACH PLANT TISSUE.
 CC -1- CATALYTIC ACTIVITY: DONOR + H₂O₂(2) = OXIDIZED DONOR + 2 H₂O.
 CC -1- COFACTOR: HEME.
 CC -1- TISSUE SPECIFICITY: ROOTS.
 CC -1- FAMILY: BELONGS TO THE PEROXIDASE FAMILY. PLANT PEROXIDASE
 CC SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M58381; AAA32842.1; -;
 CC PIR; J00458; J00458.
 CC HSP; P00433; LATJ.
 CC InterPro; IPR000823; -;
 CC DR InterPro; IPR002016; -;
 CC DR Pfam; PF00141; peroxidase; 1.
 CC DR PRINTS; PR00458; PEROXIDASE.
 CC DR PRINTS; PR00461; PLPEROXIDASE.
 CC DR PROSITE; PS00435; PEROXIDASE_1; 1.
 CC DR PROSITE; PS00436; PEROXIDASE_2; 1.
 CC Oxidoreductase; Glycoprotein; Peroxidase; Heme; Multigene family;
 CC Signal.
 CC KW SIGNAL.
 CC FT CHAIN 1 19
 CC FT ACT_SITE 20 349 BASIC PEROXIDASE E.
 CC FT ACT_SITE 67 67 BY SIMILARITY.
 CC FT ACT_SITE 71 71 DISTAL HISTIDINE (BY SIMILARITY).
 CC FT ACT_SITE 199 199 PROXIMAL HISTIDINE (HEME AXIAL LIGAND).
 CC FT DISULFID 40 120 BY SIMILARITY.
 CC FT DISULFID 73 78 BY SIMILARITY.
 CC FT DISULFID 126 329 BY SIMILARITY.
 CC FT DISULFID 206 238 BY SIMILARITY.
 CC FT CARBOHYD 86 86 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 173 173 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 187 187 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 217 217 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 243 243 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC SQ SEQUENCE 349 AA; 38172 MW; 6E02C5758C36AB8F CRC64;
 CC -----
 CC Query Match 78.0%; Score 32; DB 1; Length 349;
 CC Best Local Similarity 83.3%; Pred. No. 26;
 CC Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 CC -----
 CC QY 2 MYPTYL 7
 CC Db 226 LYPTYL 231
 CC -----
 CC RESULT 5
 CC NUSM_HIPAM STANDARD; PRT; 594 AA.
 CC AC O922Y1:
 CC DT 30-MAY-2000 (Rel. 39, Created)
 CC DT 30-MAY-2000 (Rel. 39, Last sequence update)
 CC DE NADH+UBIQUINONE OXIDOREDUCTASE (Rel. 39, Last annotation update)
 CC DE NADH+UBIQUINONE OXIDOREDUCTASE CHAIN 5 (EC 1.6.5.3).
 CC GN MTND5 OR ND5 OR NADH5.
 CC OS Hippopotamus amphibius (Hippopotamus).
 CC OC Mitochondrion.
 CC OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Cetartiodactyla; Hippopotamidae; Hippopotamus.
 CC OX NCBI_TaxID=9833;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.

Ursing B.M., Arnason U.;
 RT "Analyses of mitochondrial genomes strongly support a hippopotamus-
 RT whale clade.";
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AJ010957; CAA09438.1; -;
 CC InterPro; IPR001516; -;
 CC DR InterPro; IPR001750; -;
 CC DR Pfam; PF00361; oxidored_ql; 1.
 CC DR Pfam; PF00662; oxidored_ql_N; 1.
 CC KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
 CC SQ SEQUENCE 594 AA; 66599 MW; 4130499096B5A5CE CRC64;
 CC -----
 CC Query Match 78.0%; Score 32; DB 1; Length 594;
 CC Best Local Similarity 83.3%; Pred. No. 44;
 CC Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 CC -----
 CC QY 3 YPTVLK 8
 CC Db 32 YPTVVK 37
 CC -----
 CC RESULT 6
 CC RPOB_THEMA STANDARD; PRT; 1263 AA.
 CC AC P29398:
 CC DT 01-DEC-1992 (Rel. 24, Created)
 CC DT 01-JUN-1994 (Rel. 29, Last sequence update)
 CC DT 30-MAY-2000 (Rel. 39, Last annotation update)
 CC DE DNA-DIRECTED RNA POLYMERASE BETA CHAIN (EC 2.7.7.6) (TRANSCRIPTASE
 CC DE BETA CHAIN) RNA POLYMERASE BETA SUBUNIT).
 CC GN RPOB OR TM0458.
 CC OS Thermotoga maritima.
 CC OC Bacteria; Thermotogales; Thermotoga.
 CC OX NCBI_TaxID=2336;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=MSB8 / DSM 3109;
 CC RX MEDLINE=94232816; PubMed=8177738;
 CC RA Palm P., Schleper C., Arnold-Ammer I., Holz I., Meier T.,
 CC RA Lottspeich F., Zillig W.;
 CC RT "The DNA-dependent RNA-polymerase of Thermotoga maritima;
 CC RT characterisation of the enzyme and the DNA-sequence of the genes for
 CC RT the large subunits.";
 CC RL Nucleic Acids Res. 21:4904-4908(1993).
 CC RN [2]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=MSB8 / DSM 3109;
 CC RX MEDLINE=99287316; PubMed=10360571;
 CC RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
 CC RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
 CC RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
 CC RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
 CC RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
 CC RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
 CC RT "Evidence for lateral gene transfer between Archaea and Bacteria from
 CC RT genome sequence of Thermotoga maritima.";
 CC RL Nature 399:323-329(1999).
 CC RN [3]
 CC RP SEQUENCE OF 1-404 FROM N.A.
 CC RC STRAIN=MSB8 / DSM 3109;
 CC RX MEDLINE=93054590; PubMed=1429627;
 CC RN [1]
 CC RP Liao D., Dennis P.P.;

RT "The organization and expression of essential transcription
 RT translation component genes in the extremely thermophilic eubacterium
 RL J. Biol. Chem. 267:22787-22797(1992).
 CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
 CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
 CC SUBSTRATES.
 CC -1- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE - N PYROPHOSPHATE +
 CC RNA(N).
 CC -1- SUBUNIT: THE ENZYME CONSISTS OF THE SIGMA CHAIN AND THE CORE
 CC ENZYME WHICH IS COMPOSED OF 2 ALPHA CHAINS, 1 BETA CHAIN, AND 1
 CC BETA' CHAIN.
 CC -1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA CHAIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X72695; CAA51246.1; -.
 CC EMBL: AE001724; AAD35543.1; -.
 CC EMBL: Z11839; CAA77863.1; -.
 CC PIR: S19903; S19903.
 CC PIR: S41466; S41466.
 CC TIGR: TM0458; -.
 CC InterPro: IPR001572; -.
 CC Pfam: PF00562; RNA_POL_B; 1.
 CC PROSITE: PS01166; RNA_POL_BETA; 1.
 CC Transferrase; Transcription; DNA-directed RNA polymerase.
 CC SEQUENCE 1263 AA; 143137 MW; 04B79368567C8237 CRC64;

Query Match 78.0%; Score 32; DB 1; Length 1263;
 Best Local Similarity 83.3%; Pred. No. 94;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYL 7
 Db 214 LYPTYL 219
 RESULT 7
 ID MM07_FELCA STANDARD; PRT; 262 AA.
 AC P55032;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE MATRILYSIN PRECURSOR (EC 3.4.24.23) (PUMP-1 PROTEASE) (UTERINE
 DE METALLOPROTEINASE) (MATRIX METALLOPROTEINASE-7) (MMP-7) (MATRIN)
 DE (FRAGMENT).
 GN MMP7
 OS Felis silvestris catus (Cat).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
 CC NCBI_TaxID=9685;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC TISSUE=Endometrium;
 RA Scalzo C.M., Verhage H.G., Jaffe R.C.;
 RT "Expression and estrogen control of PUMP-1 mRNA in the cat uterus";
 RL Endocrinol. Jpn. 2:229-235(1994).
 CC -1- FUNCTION: DEGRADATES CASEIN, GELATINS OF TYPES I, III, IV, AND V,
 CC AND FIBRONECTIN. ACTIVATES PROCOLLAGENASE.
 CC -1- CATALYTIC ACTIVITY: CLEAVAGE OF 14-ALA-1-LEU-15 AND 16-TYR-1-
 CC LEU-17 IN B CHAIN OF INSULIN. NO ACTION ON COLLAGEN TYPES I, II,
 CC IV, V. CLEAVES GELATIN CHAIN ALPHA-2(1) > ALPHA-1(1).
 CC COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC
 CC METALLOPROTEASE) ALSO KNOWN AS MATRININ SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: U04444; AAA18222.1; -.
 CC HSP: P09237; IMMR.
 CC MEROPS: M10.008; -.
 CC InterPro: IPR000130; -.
 CC InterPro: IPR001818; -.
 CC Pfam: PF00413; Peptidase_M10; 1.
 CC PROSITE: PS00142; ZINC_PROTEASE; 1.
 CC PROSITE: PS00546; ZINCSTEIN_SWITCH; FALSE_NEG.
 CC Hydrolase; Metalloprotease; Zinc; Zymogen; Calcium;
 CC Collagen degradation; Extracellular matrix; Signal.
 CC NON_TER 1
 CC SIGNAL <1 12 BY SIMILARITY.
 CC FT PROPEP 13 89 ACTIVATION PEPTIDE (BY SIMILARITY).
 CC FT CHAIN 90 262 MATRILYSIN.
 CC FT SITE 82 82 CYSTEINE SWITCH (POTENTIAL).
 CC FT METAL 209 209 ZINC (CATALYTIC) (BY SIMILARITY).
 CC FT ACT_SITE 210 210 BY SIMILARITY.
 CC FT METAL 213 213 ZINC (CATALYTIC) (BY SIMILARITY).
 CC FT METAL 219 219 ZINC (CATALYTIC) (BY SIMILARITY).
 CC SQ SEQUENCE 262 AA; 29263 MW; EA1FA23320DC732 CRC64;

Query Match 75.6%; Score 31; DB 1; Length 262;
 Best Local Similarity 100.0%; Pred. No. 31;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
 Db 227 MYPTY 231
 RESULT 8
 ID MM07_MOUSE STANDARD; PRT; 264 AA.
 AC Q10738;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE MATRILYSIN PRECURSOR (EC 3.4.24.23) (PUMP-1 PROTEASE) (UTERINE
 DE METALLOPROTEINASE) (MATRIX METALLOPROTEINASE-7) (MMP-7) (MATRIN).
 GN MMP7
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CC NCBI_TaxID=10090;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC STRAIN=ICR; TISSUE=Uterus;
 CC MEDLINE=96086498; PubMed=7579699;
 CC Wilson C.L., Heppner K.J., Rudolph L.A., Matrisian L.M.;
 RT "The metalloproteinase matrilysin is preferentially expressed by
 RT epithelial cells in a tissue-restricted pattern in the mouse";
 RL Mol. Biol. Cell 6:851-869(1995).
 CC -1- FUNCTION: DEGRADATES CASEIN, GELATINS OF TYPES I, III, IV, AND V,
 CC AND FIBRONECTIN. ACTIVATES PROCOLLAGENASE (BY SIMILARITY).
 CC -1- FUNCTION: MAY PLAY A ROLE IN TISSUE REORGANIZATION.
 CC -1- CATALYTIC ACTIVITY: CLEAVAGE OF 14-ALA-1-LEU-15 AND 16-TYR-1-
 CC LEU-17 IN B CHAIN OF INSULIN. NO ACTION ON COLLAGEN TYPES I, II,
 CC IV, V. CLEAVES GELATIN CHAIN ALPHA-2(1) > ALPHA-1(1).
 CC COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC
 CC METALLOPROTEASE) ALSO KNOWN AS MATRININ SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; L36238; AAA99984.1; -;
 DR EMBL; L36243; AAA99984.1; JOINED.
 DR EMBL; L36242; AAA99984.1; JOINED.
 DR EMBL; L36241; AAA99984.1; JOINED.
 DR EMBL; L36240; AAA99984.1; JOINED.
 DR EMBL; L36239; AAA99984.1; JOINED.
 DR EMBL; L36244; AAA99983.1; -;
 DR HSSP; P09237; IMMR.
 DR MEROPS; M10.008; -;
 DR MGD; MGI:103189; Mmp7.
 DR InterPro; IPR000130; -;
 DR InterPro; IPR001818; -;
 DR Pfam; PF00413; Peptidase_M10; 1.
 DR PRINTS; PR00138; MATRIXIN.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 DR PROSITE; PS00546; CYSTEINE_SWITCH; 1.
 KW Hydrolase; Metalloprotease; Zinc; Zymogen; Calcium;
 KW Collagen degradation; Extracellular matrix; Signal.
 FT SIGNAL 1 17 POTENTIAL.
 FT PROPEP 18 94 ACTIVATION PEPTIDE (BY SIMILARITY).
 FT CHAIN 95 264 MATRILYSIN.
 FT SITE 87 87 CYSTEINE SWITCH (POTENTIAL).
 FT METAL 214 214 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACT_SITE 215 215 BY SIMILARITY.
 FT METAL 218 218 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 224 224 ZINC (CATALYTIC) (BY SIMILARITY).
 FT CONFLICT 201 201 G -> D (IN AAA99983).
 SQ SEQUENCE 264 AA; 29755 MW; EDA31A5EBAC63342 CRC64;

Query Match 75.6%; Score 31; DB 1; Length 264;
 Best Local Similarity 100.0%; Pred. No. 31;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPT 6
 DB 232 MYPT 236

RESULT 9
 MM07_HUMAN STANDARD; PRT; 267 AA.
 P09237;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE MATRILYSIN PRECURSOR (EC 3.4.24.23) (PUMP-1 PROTEASE) (UTERINE
 DE METALLOPROTEINASE) (MATRIX METALLOPROTEINASE-7) (MMP-7) (MATRIN).
 GN MMP7 OR MPSL1 OR PUMP1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88339885; PubMed=2844164;
 RA Muller D., Quantin B., Gesnel M.-C., Millon-Collard R., Abecassis J.,
 RA Breathnach R.;
 RT "The collagenase gene family in humans consists of at least four
 RT members.";
 RL Biochem. J. 253:187-192(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=92359961; PubMed=1497627;
 RA Marti H.P., McNeill L., Thomas G., Davies M., Lovett D.H.;

RT "Molecular characterization of a low-molecular-mass matrix
 RT metalloproteinase secreted by glomerular mesangial cells as PUMP-1.";
 RL Biochem. J. 285:899-905(1992).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=94124555; PubMed=8294454;
 RA Gaire M., Magbanua Z., McDonnell S., McNeill L.B., Lovett D.H.,
 RA Matrisian L.M.;
 RT "Structure and expression of the human gene for the matrix
 RT metalloproteinase matrilysin.";
 RL J. Biol. Chem. 269:2032-2040(1994).
 RN [4]
 RP SEQUENCE OF 18-42.
 RX MEDLINE=91070531; PubMed=2253219;
 RA Miyazaki K., Hattori Y., Umenishi F., Yasumitsu H., Umeda M.;
 RT "Purification and characterization of extracellular matrix-degrading
 RT metalloproteinase, matrin (pump-1), secreted from human rectal
 RT carcinoma cell line.";
 RL Cancer Res. 50:7758-7764(1990).
 RN [5]
 RP FUNCTION.
 RX MEDLINE=89375247; PubMed=2550050;
 RA Quantin B., Murphy G., Breathnach R.;
 RT "pump-1 cDNA codes for a protein with characteristics similar to
 RT those of classical collagenase family members.";
 RL Biochemistry 28:5327-5334(1989).
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
 RX MEDLINE=9575856; PubMed=7756291;
 RA Browner M.F., Smith W.W., Castelhan A.L.;
 RT "Matrilysin-inhibitor complexes: common themes among
 RT metalloproteases.";
 RL Biochemistry 34:6602-6610(1995).
 CC -!- FUNCTION: DEGRADATES CASEIN, GELATIN OF TYPES I, III, IV, AND V,
 CC AND FIBRONECTIN. ACTIVATES PROCOLLAGENASE.
 CC -!- CATALYTIC ACTIVITY: CLEAVAGE OF 14-ALA-1-LEU-15 AND 16-TYR-1-
 CC LEU-17 IN B CHAIN OF INSULIN. NO ACTION ON COLLAGEN TYPES I, II,
 CC IV, V. CLEAVES GELATIN CHAIN ALPHA-2(I) > ALPHA-1(I).
 CC -!- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC
 CC METALLOPROTEASE) ALSO KNOWN AS MATRININ SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; X07819; CAA30678.1; -;
 DR EMBL; Z11887; CAA77942.1; -;
 DR EMBL; L22524; AAC37543.1;
 DR EMBL; L22519; AAC37543.1; JOINED.
 DR EMBL; L22520; AAC37543.1; JOINED.
 DR EMBL; L22521; AAC37543.1; JOINED.
 DR EMBL; L22522; AAC37543.1; JOINED.
 DR EMBL; L22523; AAC37543.1; JOINED.
 DR PIR; B28816; KCHUM.
 DR PIR; S24324; S24324.
 DR PDB; 1MMQ; 03-APR-96.
 DR PDB; 1MMR; 03-APR-96.
 DR MEROPS; M10.008; -;
 DR MIM; I78990; -;
 DR InterPro; IPR000130; -;
 DR InterPro; IPR001818; -;
 DR Pfam; PF00413; Peptidase_M10; 1.
 DR PRINTS; PR00138; MATRIXIN.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 DR PROSITE; PS00546; CYSTEINE_SWITCH; 1.
 KW Hydrolase; Metalloprotease; Zinc; Zymogen; Calcium; 3D-structure;

KW Collagen degradation; Extracellular matrix; Signal; Polymorphism.
 FT SIGNAL 1 17
 FT PROPEP 18 94
 FT CHAIN 95 267
 FT SITE 87 87
 FT METAL 214 214
 FT ACT SITE 215 215
 FT METAL 218 218
 FT METAL 224 224
 FT VARIANT 77 77
 R -> H.
 /FTID=VAR_006729;
 SQ SEQUENCE 267 AA; 29677 MW; F6BDFC0ADAZ3603 CRC64;

Query Match 75.6%; Score 31; DB 1; Length 267;
 Best Local Similarity 100.0%; Pred. No. 32;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
 |||||
 232 MYPTY 236

RESULT 10
 ID MM07_RAT STANDARD; PRT; 267 AA.
 AC P50280;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE MATRILYSIN PRECURSOR (EC 3.4.24.23) (PUMP-1
 DE METALLOPROTEINASE) (MATRIX METALLOPROTEINASE-7) (MMP-7) (MATRIN).
 GN MMP7 OR MMP-7.
 OS Rattus norvegicus (Rat).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Uterus;
 RX MEDLINE=95332299; PubMed=7608162;
 RA Abramson S.R., Conner G.E., Nagase H., Neuhaus I., Woessner J.F.;
 RT "Characterization of rat uterine matrilysin and its cDNA.
 RT Relationship to human pump-1 and activation of procollagenases.";
 RL J. Biol. Chem. 270:16016-16022(1995).
 CC -1- FUNCTION: DEGRADATES CASEIN, GELATINS OF TYPES I, III, IV, AND V,
 CC AND FIBRONECTIN. ACTIVATES PROCOLLAGENASE (BY SIMILARITY).
 CC LEU-17 IN B CHAIN OF INSULIN. NO ACTION ON COLLAGEN TYPES I, II,
 CC IV, V. CLEAVES GELATIN CHAIN ALPHA-2(I) > ALPHA-1(I).
 CC -1- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC
 CC METALLOPROTEASE) ALSO KNOWN AS MATRIN SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L24374; AAA99432.1;
 CC HSP; P09237; IMR.
 CC MEROPS; M10.008;
 CC InterPro; IPR000130;
 CC InterPro; IPR001818;
 CC Pfam; PF00413; Peptidase_M10; 1.
 CC PRINTS; PR00138; MATRIN.
 CC PROSITE; PS00142; ZINC_PROPEASE; 1.
 CC PROSITE; PS00546; CYSTEINE_SWITCH; 1.
 KW Hydrolase; Metalloprotease; Zinc; Zymogen; Calcium;
 KW Collagen degradation; Extracellular matrix; Signal.

FT SIGNAL 1 20
 FT PROPEP 21 97
 FT CHAIN 98 267
 FT SITE 90 90
 FT METAL 217 217
 FT ACT SITE 218 218
 FT METAL 221 221
 FT METAL 227 227
 FT METAL 232 232
 SQ SEQUENCE 267 AA; 29885 MW; EBA3C3D9527A4C7B CRC64;

Query Match 75.6%; Score 31; DB 1; Length 267;
 Best Local Similarity 100.0%; Pred. No. 32;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
 |||||
 235 MYPTY 239

RESULT 11
 ID SC14_CANAL STANDARD; PRT; 301 AA.
 AC P46250;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 30-NOV-2000 (Rel. 39, Last annotation update)
 DE SEC14 CYTOSOLIC FACTOR (PHOSPHATIDYLINOSITOL/PHOSPHATIDYLCHOLINE
 DE TRANSFER PROTEIN) (PI/PC TP).
 GN SEC14.
 OS Candida albicans (Yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; mitosporic Saccharomycetales; Candida.
 OX NCBI_TaxID=5476;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SGY243;
 RX MEDLINE=98048477; PubMed=9387231;
 RA Riggle P.J., Slobodkin I.V., Brown D.H. Jr., Hanson M.P.,
 RA Volkert T.L., Kumamoto C.A.;
 RT "Two transcripts, differing at their 3' ends, are produced from the
 RT Candida albicans SEC14 gene."
 RL Microbiology 143:3527-3535(1997).
 CC -1- FUNCTION: REQUIRED FOR TRANSPORT OF SECRETORY PROTEINS FROM THE
 CC GOLGI COMPLEX. CATALYZES THE TRANSFER OF PHOSPHATIDYLINOSITOL AND
 CC PHOSPHATIDYLCHOLINE BETWEEN MEMBRANES IN VITRO (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: ASSOCIATED WITH THE GOLGI COMPLEX AS A
 CC PERIPHERAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SEC14 CYTOSOLIC FACTOR FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X81937; CAA57490.1;
 CC HSP; U61975; AAB41491.1;
 CC HSP; P24280; IAU.
 CC InterPro; IPR001251;
 CC Pfam; PF00650; CRAL_TRIO; 1.
 KW Transport; Protein transport; Golgi stack
 KW SEQUENCE 301 AA; 34709 MW; 1F1948EA8B525BE CRC64;

Query Match 75.6%; Score 31; DB 1; Length 301;
 Best Local Similarity 100.0%; Pred. No. 36;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
 |||||
 Db 106 MYPTY 110

RESULT 12
 MM12_MOUSE STANDARD; PRT; 462 AA.
 AC P34960;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE MACROPHAGE METALLOELASTASE PRECURSOR (EC 3.4.24.65) (MME) (MATRIX METALLOPROTEINASE-12) (MMP-12).
 DE MMP12 OR MMEL OR MME.
 Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 99-125.
 RC TISSUE=Macrophage;
 RX MEDLINE=92165826; PubMed=1537850;
 RA Shapiro S.D., Griffin G.L., Gilbert D.J., Jenkins N.A.,
 RA Copeland N.G., Welgus H.G., Senior R.M., Ley T.J.,
 RT "Molecular cloning, chromosomal localization, and bacterial expression of a murine macrophage metalloelastase.";
 RT J. Biol. Chem. 267:4664-4671(1992).
 CC -!- FUNCTION: MAY BE INVOLVED IN TISSUE INJURY AND REMODELING. HAS
 CC SIGNIFICANT ELASTOLYTIC ACTIVITY.
 CC -!- CATALYTIC ACTIVITY: HYDROLYSIS OF SOLUBLE AND INSOLUBLE ELASTIN.
 CC SPECIFIC CLEAVAGES ARE ALSO PRODUCED AT 14-ALA-|-LEU-15 AND 16-
 CC TYR-|-LEU-17 IN THE B CHAIN OF INSULIN.
 CC -!- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY.
 CC -!- SIMILARITY: CONTAINS 1 HEMOPEXIN-LIKE DOMAIN.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC
 CC METALLOPROTEASE) ALSO KNOWN AS MATRININ SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M82831; AAA39526.1; -
 CC PIR; A42401; A42401.
 CC HSSP; P03956; ICGL.
 CC MEROPS; M10.009; -
 CC MGD; MGI:97005; Mmp12.
 CC InterPro; IPR000130; -
 CC InterPro; IPR000585; -
 CC InterPro; IPR001818; -
 CC Pfam; PF000413; Peptidase_M10; 1.
 CC Pfam; PF00045; hemopexin; 4.
 CC PRINTS; PR00138; MATRININ.
 CC PROSITE; PS00024; ZINC_PROTEASE; 1.
 CC PROSITE; PS00142; ZINC_PROTEASE; 1.
 CC PROSITE; PS00546; CYSTEINE_SWITCH; 1.
 CC Hydrolase; Metalloprotease; Glycoprotein; Zinc; Zymogen; Calcium;
 CC Extracellular matrix; Signal.
 CC SIGNAL 1 17 PROBABLE.
 FT PROPP 18 98 ACTIVATION PEPTIDE.
 FT CHAIN 99 462 MACROPHAGE METALLOELASTASE.
 FT DOMAIN 272 462 HEMOPEXIN-LIKE.
 FT SITE 85 85 CYSTEINE SWITCH (BY SIMILARITY).

FT METAL 211 211 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACT_SITE 212 212 BY SIMILARITY.
 FT METAL 215 215 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 221 221 ZINC (CATALYTIC) (BY SIMILARITY).
 FT CARBOHYD 21 21 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 74 74 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 310 310 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT DISULFID 275 462 BY SIMILARITY.
 SQ SEQUENCE 462 AA; 53841 MW; BB9625906FIDBEDF CRC64;
 Query Match 75.6%; Score 31; DB 1; Length 462;
 Best Local Similarity 100.0%; Pred. No. 55;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 MYPTY 6
 |||||
 Db 229 MYPTY 233
 RESULT 13
 MM12_RAT STANDARD; PRT; 465 AA.
 ID MM12_RAT
 AC Q63341;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE MACROPHAGE METALLOELASTASE PRECURSOR (EC 3.4.24.65) (MME) (MATRIX METALLOPROTEINASE-12) (MMP-12).
 DE MMP12 OR MMEL.
 GN Rattus norvegicus (Rat).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY;
 RA Cossins J., Clements J., Catlin G.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: MAY BE INVOLVED IN TISSUE INJURY AND REMODELING. HAS
 CC SIGNIFICANT ELASTOLYTIC ACTIVITY (BY SIMILARITY).
 CC -!- CATALYTIC ACTIVITY: HYDROLYSIS OF SOLUBLE AND INSOLUBLE ELASTIN.
 CC SPECIFIC CLEAVAGES ARE ALSO PRODUCED AT 14-ALA-|-LEU-15 AND 16-
 CC TYR-|-LEU-17 IN THE B CHAIN OF INSULIN.
 CC -!- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY (BY SIMILARITY).
 CC -!- SIMILARITY: CONTAINS 1 HEMOPEXIN-LIKE DOMAIN.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC
 CC METALLOPROTEASE) ALSO KNOWN AS MATRININ SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X98517; CAA67142.1; -
 CC HSSP; P03956; ICGL.
 CC MEROPS; M10.009; -
 CC InterPro; IPR000130; -
 CC InterPro; IPR000585; -
 CC InterPro; IPR001818; -
 CC Pfam; PF000413; Peptidase_M10; 1.
 CC Pfam; PF00045; hemopexin; 4.
 CC PRINTS; PR00138; MATRININ.
 CC PROSITE; PS00024; ZINC_PROTEASE; 1.
 CC PROSITE; PS00142; ZINC_PROTEASE; 1.
 CC PROSITE; PS00546; CYSTEINE_SWITCH; 1.
 CC Hydrolase; Metalloprotease; Glycoprotein; Zinc; Zymogen; Calcium;
 CC Extracellular matrix; Signal.
 CC SIGNAL 1 21 PROBABLE.
 FT PROPP 22 101 ACTIVATION PEPTIDE (BY SIMILARITY).
 FT SITE 21 21

```

FT CHAIN 102 465 MACROPHAGE METALLOELASTASE.
FT DOMAIN 275 465 HEMOXIN-LIKE.
FT SITE 88 214 CYSTEINE SWITCH (BY SIMILARITY).
FT METAL 214 214 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 215 215 BY SIMILARITY.
FT METAL 218 218 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 224 224 ZINC (CATALYTIC) (BY SIMILARITY).
FT CARBOHYD 313 313 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 278 465 BY SIMILARITY.
SQ SEQUENCE 465 AA; 53738 MW; E779B6014EC6FF68 CRC64;

Query Match 75.6%; Score 31; DB 1; Length 465;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
DB 232 MYPTY 236

RESULT 14
MM18_XENLA STANDARD; PRT; 467 AA.
ID MM18_XENLA
AC OL3065;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DE MATRIX METALLOPROTEINASE-18 PRECURSOR (EC 3.4.24.-) (MMP-18)
DE (COLLAGENASE-4) (COLLAGENASE 4) (XCOLA).
GN MMP18 OR COL4.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoides; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Intestine;
RX MEDLINE=97053976; PubMed=8989355;
RA Stelow M.A., Baizon D.D., Li J., Sedgwick T., Liang V.C.-T.,
RA Sang Q.A., Shi Y.-B.;
RT Identification and characterization of a novel collagenase in Xenopus
RL laevis: possible roles during frog development.";
RM Mol. Biol. Cell 7:1471-1483(1996).
CC -!- FUNCTION: CLEAVES COLLAGEN TYPE I. MAY PLAY A ROLE IN LARVAL
CC TISSUE DEGENERATION AND ADULT ORGANOGENESIS DURING AMPHIBIAN
CC METAMORPHOSIS. MAY BE INVOLVED IN TAIL RESORPTION.
CC -!- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY (BY SIMILARITY).
CC -!- ENZYME REGULATION: UPREGULATED IN THE TAIL BY THYROID HORMONE.
CC -!- SUBCELLULAR LOCATION: SECRETED. ASSOCIATED WITH THE EXTRACELLULAR
CC MATRIX (BY SIMILARITY).
CC -!- TISSUE SPECIFICITY: EXPRESSED ONLY TRANSIENTLY IN WHOLE ANIMAL, AT
CC TIME WHEN TADPOLE FEEDING BEGINS.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED AT HIGH LEVEL AS THE TADPOLE TAIL
CC RESORBS AND DURING HINDLIMB MORPHOGENESIS AND INTESTINAL
CC REMODELING.
CC -!- SIMILARITY: CONTAINS 1 HEMOXIN-LIKE DOMAIN.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC
CC METALLOPROTEASE) ALSO KNOWN AS MATRIXIN SUBFAMILY.
CC -----
CC This swiss-prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L76275; AAB53148.1;
CC HSSP: P03956; ICGI.
CC MEROPS: M10.018.
CC InterPro: IPR000130;

```

```

DR InterPro: IPR000585;
DR InterPro: IPR001818;
DR Pfam: PF00045; hemopexin: 4
DR Pfam: PF00413; peptidase_M10: 1.
DR PRINTS: PS00136; MATRIXIN.
DR PROSITE: PS00142; ZINC_PROTEASE; 1. FALSE_NEG.
DR PROSITE: PS00546; CYSTEINE_SWITCH; 1. FALSE_NEG.
DR PROSITE: PS00534; HEMOXIN; FALSE_NEG.
KW Hydrolase; Metalloprotease; Zinc; Calcium; zymogen; signal;
KW Collagen degradation; Extracellular matrix.
FT SIGNAL 1 17 POTENTIAL.
FT PROPEP 18 99 BY SIMILARITY.
FT CHAIN 100 467 MATRIX METALLOPROTEINASE-18.
FT DOMAIN 277 467 HEMOXIN-LIKE.
FT SITE 92 92 CYSTEINE SWITCH (POTENTIAL).
FT METAL 218 218 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 219 219 BY SIMILARITY.
FT METAL 222 222 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 228 228 ZINC (CATALYTIC) (BY SIMILARITY).
FT DISULFID 280 467 POTENTIAL.
SQ SEQUENCE 467 AA; 52812 MW; 4623F6CE3454051 CRC64;

Query Match 75.6%; Score 31; DB 1; Length 467;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
DB 236 MYPTY 240

RESULT 15
MM20_MOUSE STANDARD; PRT; 482 AA.
ID MM20_MOUSE
AC P57748;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE MATRIX METALLOPROTEINASE-20 PRECURSOR (EC 3.4.24.-) (MMP-20) (ENAMEL
DE METALLOPROTEINASE) (ENAMELYSIN).
GN MMP20.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SVJ;
RX MEDLINE=20079167; PubMed=10610728;
RA Caterina J., Shi J., Krakora S., Bartlett J.D., Engler J.A.,
RA Kozak C.A., Birkedal-Hansen H.;
RT Isolation, characterization, and chromosomal location of the mouse
RT enamelysin gene."
RL Genomics 62:308-311(1999).
CC -!- FUNCTION: DEGRADATES AMELOGENIN, THE MAJOR PROTEIN COMPONENT OF THE
CC ENAMEL MATRIX AND TWO OF THE MACROMOLECULES CHARACTERISING THE
CC CARTILAGE EXTRACELLULAR MATRIX: AGGREGAN AND THE CARTILAGE
CC OLIGOMERIC MATRIX PROTEIN (COMP). MAY PLAY A CENTRAL ROLE IN TOOTH
CC ENAMEL FORMATION (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: CLEAVES AGGREGAN AT THE 360-ASN-1-PHE-361 SITE
CC (BY SIMILARITY).
CC -!- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: SECRETED. ASSOCIATED WITH THE EXTRACELLULAR
CC MATRIX (BY SIMILARITY).
CC -!- TISSUE SPECIFICITY: EXPRESSED SPECIFICALLY IN THE ENAMEL ORGAN.
CC -!- PFM: AUTOACTIVATES AT LEAST AT THE 106-ASN-1-TYR-107 SITE (BY
CC SIMILARITY).
CC -!- SIMILARITY: CONTAINS 1 HEMOXIN-LIKE DOMAIN.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC
CC METALLOPROTEASE) ALSO KNOWN AS MATRIXIN SUBFAMILY.
CC -----
CC This swiss-prot entry is copyright. It is produced through a collaboration

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; AF156956; AAF28472.1; -
DR EMBL; AF156947; AAF28472.1; JOINED.
DR EMBL; AF156948; AAF28472.1; JOINED.
DR EMBL; AF156949; AAF28472.1; JOINED.
DR EMBL; AF156950; AAF28472.1; JOINED.
DR EMBL; AF156951; AAF28472.1; JOINED.
DR EMBL; AF156952; AAF28472.1; JOINED.
DR EMBL; AF156953; AAF28472.1; JOINED.
DR EMBL; AF156954; AAF28472.1; JOINED.
DR EMBL; AF156955; AAF28472.1; JOINED.
DR EMBL; AF155933; AAF28470.1; -
DR MEROPS; M10.019; -
MGD; MGI:1353466; Mmp20.
PROSITE; PS00546; CYSTEINE_SWITCH; 1.
PROSITE; PS00142; ZINC_PROTEASE; 1.
DR PROSITE; PS00024; HEMOPEXIN; FALSE_NEG.
KW Hydrolase; Metalloprotease; Zinc; Calcium; Zymogen; Signal;
KW Extracellular matrix
FT SIGNAL 1
FT PROPEP 22 106 POTENTIAL.
FT CHAIN 107 482 BY SIMILARITY.
FT DOMAIN 292 482 MATRIX METALLOPROTEINASE-20.
FT SITE 99 99 HEMOPEXIN-LIKE.
FT METAL 225 225 CYSTEINE SWITCH (POTENTIAL).
FT ACT_SITE 226 226 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 229 229 BY SIMILARITY.
FT METAL 235 235 ZINC (CATALYTIC) (BY SIMILARITY).
FT DISULFID 295 482 ZINC (CATALYTIC) (BY SIMILARITY).
SQ SEQUENCE 482 AA; 54373 MW; 366149FAF2BDC8BB CRC64;

Query Match 75.6%; Score 31; DB 1; Length 482;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPT 6
Db 243 MYPT 247
|||||

rch completed: September 4, 2001, 15:54:09
time: 306 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2001, 15:53:35 ; Search time 80.44 Seconds
(without alignments)
13.158 Million cell updates/sec

Title: US-09-630-345-2

Perfect score: 41

Sequence: 1 XMYPTYLK 8

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_16:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phase:*

10: sp_plant:*

11: sp_rodent:*

12: sp_unclassified:*

13: sp_vertebrate:*

14: sp_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35	85.4	1088	4 Q9H2T7	Q9h2t7 homo sapien
2	33	80.5	269	1 Q27159	Q27159 methanobact
3	33	80.5	1069	10 Q9L7Z8	Q9L7z8 arabidopsis
4	32	78.0	252	2 Q9L875	Q9L875 haemophilus
5	32	78.0	553	11 Q70577	Q70577 mus musculus
6	32	78.0	555	11 Q9R0W2	Q9R0w2 rattus norv
7	32	78.0	593	11 P97558	P97558 rattus norv
8	32	78.0	593	11 P70485	P70485 rattus norv
9	32	78.0	986	4 Q94858	Q94858 homo sapien
10	32	78.0	1346	4 Q9UIZ3	Q9uiz3 homo sapien
11	31	75.6	40	14 Q9IwZ8	Q9Iwz8 human coxa
12	31	75.6	40	14 Q9IwZ7	Q9Iwz7 human coxa
13	31	75.6	50	14 Q66796	Q66796 human echov
14	31	75.6	53	6 Q9XS83	Q9xs83 equus cabal
15	31	75.6	198	14 Q9Q907	Q9q907 shope fibro
16	31	75.6	261	4 Q9NRE1	Q9nrel homo sapien
17	31	75.6	261	4 Q9NR87	Q9nr87 homo sapien
18	31	75.6	261	4 Q9G2S2	Q9g2s2 homo sapien
19	31	75.6	262	5 Q19918	Q19918 caenorhabdi

Q9tv55 sus scrofa
Q94467 dictyosteli
Q9vj65 drosophila
Q51888 prevotella
Q9shh5 arabidopsis
Q9ska2 arabidopsis
Q9hd2 homo sapien
Q9vpa2 drosophila
Q98857 cynops pyrr
Q96640 bovine pap
Q18426 caenorhabdi
Q18198 caenorhabdi
Q29350 archaeoglob
Q9txnl caenorhabdi
Q9esn3 mus musculu
Q88445 swine vesic
Q9j919 human coxa
Q02173 caenorhabdi
Q91401 pseudomonas
Q98534 paramecium
Q9vuzl drosophila
Q9v5y8 drosophila
Q9qb94 yaba monkey
Q48871 lactobacilli
Q9dhm6 yaba-like d
Q9pia0 campylobact

ALIGNMENTS

RESULT 1

Q9H2T7 ID Q9H2T7 PRELIMINARY; PRT; 1088 AA.
AC Q9H2T7;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE RANBP17.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20568271; PubMed=11024021;
RA Kutay U., Hartmann E., Treichel N., Calado A., Carmo-Fonseca M.,
RA Prehn S., Kraft R., Gorlich D., Bischoff F.R.;
RT Identification of Two Novel RanGTP-binding Proteins Belonging to the
RT Importin beta Superfamily.;
RL J. Biol. Chem. 275:40163-40168(2000).
DR EMBL; AF222747; AGA4255.1; -.
SQ SEQUENCE 1088 AA; 124374 MW; A95D5599388EAEFB CRC64;

Query Match 85.4%; Score 35; DB 4; Length 1088;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYL 7
| | | | |
Db 744 MYPTYL 749

RESULT 2

Q27159 ID Q27159 PRELIMINARY; PRT; 269 AA.
AC Q27159;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE HYPOTHETICAL 31.5 KDA PROTEIN.
GN MTH1087.

OS Methanobacterium thermoautotrophicum.
 OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
 CC Methanothermobacter.
 OX NCBI_TaxID=145262;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN-DELTA H;
 RX MEDLINE=98037514; PubMed=9371463;
 RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., DuBois J.,
 RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
 RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
 RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
 RA Jiwani N., Caruso A., Bush D., Safer H., Pakel D., Prabhakar S.,
 RA McQuigall S., Shimer G., Royce P., Goyal A., Pietrowski S., Church G.M.,
 RA Daniels C.J., Mao J.-I., Rice P., Nolling J., Sieve J.N.:
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum
 RT deltaH: functional analysis and comparative genomics.";
 RL J. Bacteriol. 179:7135-7155(1997).
 DR EMBL: AF000879; AAB85576.1;
 KW Hypothetical protein.
 SQ SEQUENCE 269 AA; 4B7D64EBDC6B8EA7 CRC64;

Query Match 80.5%; Score 33; DB 1; Length 269;
 Best Local Similarity 83.3%; Pred. No. 52;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 3 YPTYLK 8
 DB 175 YPTVIK 180
 |||||

RESULT 3
 ID Q9L7Z8 PRELIMINARY; PRT; 1069 AA.
 AC Q9L7Z8
 DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DE SIMILARITY TO: NEGATIVE REGULATOR OF VESICLE FORMATION.
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 CC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 CC Brassicales; Brassicaceae; Arabidopsi.
 OX NCBI_TaxID=3702;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN-COLUMBIA;
 RX Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;
 Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 DE [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-COLUMBIA;
 RX MEDLINE=20277480; PubMed=10819329;
 RA Nakamura Y.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
 RT features of the regions of 4,504,864 bp covered by sixty P1 and TAC
 RT clones.";
 RL DNA Res. 7:131-135(2000).
 DR EMBL: AB024028; BAA95706.2;
 DR EMBL: AP000381; BAA95706.2; JOINED.
 DR InterPro: IPR000379;
 DR InterPro: IPR000734;
 DR PROSITE: PS00120; LIPASE_SER; UNKNOWN_1.
 SQ SEQUENCE 1069 AA; 118974 MW; 0BA354556BEC6B44 CRC64;

Query Match 80.5%; Score 33; DB 10; Length 1069;
 Best Local Similarity 83.3%; Pred. No. 2,3e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 2 MYPTYL 7
 |||||

Db 71 MYPTYL 76
 RESULT 4
 ID Q9L875 PRELIMINARY; PRT; 252 AA.
 AC Q9L875
 DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DE BETA 1-4 GLUCOSYLTRANSFERASE LGTF.
 GN LGTF.
 OS Haemophilus ducreyi.
 CC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 CC Haemophilus.
 OX NCBI_TaxID=730;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=35000;
 RX MEDLINE=20278119; PubMed=10816485;
 RA Fillardault M.J., Gibson B.W., Schilling B., Sun S., Munson R.S., Jr.,
 RA Campagnari A.A.;
 RT "Construction and Characterization of Haemophilus ducreyi
 RT Lipooligosaccharide (LOS) Mutants Defective in Expression of
 RT Heptosyltransferase III and Beta1,4-Glucosyltransferase;
 RT Identification of LOS Glycoterms Containing Lactosamine Repeats.";
 RL Infect. Immun. 68:3352-3361(2000).
 DR EMBL: AF215936; AAF72876.1;
 DR InterPro: IPR001173;
 DR Pfam: PF00535; Glycos_transf_2; 1.
 KW Transferase.
 SQ SEQUENCE 252 AA; 28736 MW; 20297CB79C070AD4 CRC64;
 Query Match 78.0%; Score 32; DB 2; Length 252;
 Best Local Similarity 83.3%; Pred. No. 78;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 2 MYPTYL 7
 DB 134 LYPTYL 139
 |||||
 RESULT 5
 ID Q70577 PRELIMINARY; PRT; 553 AA.
 AC Q70577
 DT 01-AUG-1998 (TREMblrel. 07, Created)
 DT 01-AUG-1998 (TREMblrel. 07, Last sequence update)
 DE ORGANIC CATION TRANSPORTER 2.
 GN SLC22A2 OR OCT2.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=129; TISSUE=LIVER;
 RA Mooslehner K.A.;
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
 DR EMBL: AJ006036; CAA06827.1;
 DR MGD: MGI:1335072; SLC22a2.
 DR InterPro: IPR001066;
 DR Pfam: PF00083; sugtr; 1.
 DR PROSITE: PS00216; SUGAR_TRANSPORT_1; UNKNOWN_2.
 KW Transmembrane.
 SQ SEQUENCE 553 AA; 61830 MW; 1E9744F0D5415483 CRC64;
 Query Match 78.0%; Score 32; DB 11; Length 553;

Best Local Similarity 57.1%; Pred. No. 1.8e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 4; Conservative 3;

Qy 2 MYPTLK 8

Db 457 LYPTIR 463

RESULT 6

Q9ROW2 ID Q9ROW2 PRELIMINARY; PRT; 555 AA.
AC Q9ROW2
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE ORGANIC CATION TRANSPORTER OCT2R.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
[1]
SEQUENCE FROM N.A.
RC TISSUE=KIDNEY;
RX MEDLINE=99316020; PubMed=10385678;
RA Gruendemann D., Liebig G., Kiefer N., Koester S., Schoemig E.;
RT "Selective substrates for non-neuronal monoamine transporters.";
RL Mol. Pharmacol. 56:1-10(1999).
DR EMBL; Y13154; CAB52215.1; -;
DR InterPro; IPR001066; -;
DR Pfam; PF00083; sugar_tr; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_2.
SQ SEQUENCE 555 AA; 62342 MW; 29521969AEIAC206 CRC64;

Query Match 78.0%; Score 32; DB 11; Length 555;

Best Local Similarity 57.1%; Pred. No. 1.8e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 4; Conservative 3;

Qy 2 MYPTLK 8

Db 457 LYPTIR 463

RESULT 7

P97558 ID P97558 PRELIMINARY; PRT; 593 AA.
AC P97558
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DE ORGANIC CATION TRANSPORTER.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
[1]
SEQUENCE FROM N.A.
RC TISSUE=KIDNEY;
RA Gorboulev V.G., Koepsell H.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
DR EMBL; X98334; CAA66979.1; -;
DR InterPro; IPR001066; -;
DR Pfam; PF00083; sugar_tr; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_2.
KW Ion transport; Transmembrane.
SQ SEQUENCE 593 AA; 66100 MW; 36C3E1B5DC057790 CRC64;

Query Match 78.0%; Score 32; DB 11; Length 593;

Best Local Similarity 57.1%; Pred. No. 2e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 4; Conservative 3;

Qy 2 MYPTLK 8

Db 457 LYPTIR 463

RESULT 8

P70485 ID P70485 PRELIMINARY; PRT; 593 AA.
AC P70485
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE ORGANIC CATION TRANSPORTER OCT2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
[1]
SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=KIDNEY;
RX MEDLINE=96295517; PubMed=8702418;
RA Okuda M., Saito H., Urakami Y., Takano M., Inui K.;
RT "cDNA cloning and functional expression of a novel rat kidney organic
RT cation transporter, OCT2.";
RL Biochem. Biophys. Res. Commun. 224:500-507(1996).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
DR EMBL; D83044; BAA11754.1; -;
DR InterPro; IPR001066; -;
DR Pfam; PF00083; sugar_tr; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_2.
KW Transmembrane.
SQ SEQUENCE 593 AA; 66080 MW; 36C1044E0C04B443 CRC64;

Query Match 78.0%; Score 32; DB 11; Length 593;

Best Local Similarity 57.1%; Pred. No. 2e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 4; Conservative 3;

Qy 2 MYPTLK 8

Db 457 LYPTIR 463

RESULT 9

O94858 ID O94858 PRELIMINARY; PRT; 986 AA.
AC O94858
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE KIAA0758 PROTEIN (FRAGMENT).
GN KIAA0758.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
RC TISSUE=BRIN;
RX MEDLINE=99087487; PubMed=9872452;
RA Nagase T., Ishikawa K., Suyama M., Kikuno R., Miyajima N., Tanaka A.,
RA Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XI.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 5:277-286(1998).
DR EMBL; AB018301; BAA34478.1; -;
DR InterPro; IPR000203; -;
DR Pfam; PF00002; 7tm_2; 1.
DR PROSITE; PS00002; 7tm_2; 1.
SQ SEQUENCE 986 AA; 10600 MW; 36C1044E0C04B443 CRC64;

DR PRINTS; PR00249; GPCRSECRETIN.
 DR SMART; SM00303; GFS; 1.
 FT NON_TER 1
 SQ SEQUENCE 986 AA; 108720 MW; EF4875B03BAA0E7D CRC64;

Query Match 78.0%; Score 32; DB 4; Length 986;
 Best Local Similarity 85.7%; Pred. No. 3.4e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 MYPTYLK 8
 Db 384 MLPTYLK 390
 :|||||

RESULT 10

Q9UIZ3 PRELIMINARY; PRT; 1346 AA.
 AC Q9UIZ3
 DT 01-MAY-2000 (TRENBLrel. 13, Created)
 DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
 DE VP2 PROTEIN (FRAGMENT).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OX Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]_TaxID=9606;
 RP SEQUENCE FROM N.A.
 RA Serrhini;
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX DOMAINS.
 DR EMBL: AL096772; CAB61578.1;
 DR InterPro: IPR000082;
 DR InterPro: IPR000203;
 DR InterPro: IPR000561;
 DR InterPro: IPR003082;
 DR InterPro: IPR003086;
 DR Pfam: PF00002; 7um_2; 1.
 DR Pfam: PF00047; 4um_2; 1.
 DR Pfam: PF01825; GFS; 1.
 DR PRINTS; PR00249; GPCRSECRETIN.
 DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
 DR SMART; SM00200; SEA; 1.
 SQ SEQUENCE 1346 AA; 149456 MW; 72A9D02B08218A60 CRC64;

Query Match 78.0%; Score 32; DB 4; Length 1346;
 Best Local Similarity 85.7%; Pred. No. 4.7e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 MYPTYLK 8
 Db 744 MLPTYLK 750
 :|||||

RESULT 11

Q9IWZ8 PRELIMINARY; PRT; 40 AA.
 AC Q9IWZ8
 DT 01-OCT-2000 (TRENBLrel. 15, Created)
 DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
 DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)
 DE VP2 PROTEIN (FRAGMENT).
 OS Human coxsackievirus B5.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OX Enterovirus.
 OX NCBI_TaxID=12074;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN-R8597;
 RA Casas I., Palacios G.F., Trallero G., Cisterna D., Freire M.C.,
 RT "Molecular characterization of human enteroviruses in clinical samples by three different RT nested PCR assays and direct sequencing of amplified products."
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF252193; AAF71330.1;
 DR InterPro: IPR001676;
 DR Pfam: PF00073; rhv; 1.
 DR NON_TER 1
 FT NON_TER 40
 SQ SEQUENCE 40 AA; 4388 MW; 7F998A9865EF9F67 CRC64;

Query Match 75.6%; Score 31; DB 14; Length 40;
 Best Local Similarity 71.4%; Pred. No. 17;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 MYPTYLK 8
 Db 26 VWPTYLK 32
 :|||||

RESULT 12

Q9IWZ7 PRELIMINARY; PRT; 40 AA.
 AC Q9IWZ7
 DT 01-OCT-2000 (TRENBLrel. 15, Created)
 DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
 DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)
 DE VP2 PROTEIN (FRAGMENT).
 OS Human coxsackievirus B5.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OX Enterovirus.
 OX NCBI_TaxID=12074;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-R8497;
 RA Casas I., Palacios G.F., Trallero G., Cisterna D., Freire M.C.,
 RT "Molecular characterization of human enteroviruses in clinical samples by three different RT nested PCR assays and direct sequencing of amplified products."
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF252194; AAF71331.1;
 DR InterPro: IPR001676;
 DR Pfam: PF00073; rhv; 1.
 DR NON_TER 1
 FT NON_TER 40
 SQ SEQUENCE 40 AA; 4388 MW; 7F998A9865EF9F67 CRC64;

Query Match 75.6%; Score 31; DB 14; Length 40;
 Best Local Similarity 71.4%; Pred. No. 17;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 MYPTYLK 8
 Db 26 VWPTYLK 32
 :|||||

RESULT 13

Q66796 PRELIMINARY; PRT; 50 AA.
 AC Q66796
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)
 DE CAPSID PROTEIN VP2 (FRAGMENT).
 OS Human echovirus 20.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCBI_TaxID=47508;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-JV-1;
RX MEDLINE=96203959; PubMed=8627260;
RA Huttunen P., Santti J., Pulli T., Hyypia T.;
RT "The major echovirus group is genetically coherent and related to
RT coxsackie B viruses."
RL J. Gen. Virol. 77:715-725(1996).
DR EMBL: X89549; CAA61727.1; -
DR InterPro: IPR001676; -
DR Pfam: PF00073; rnv; 1.
FT NON_TER 1
FT NON_TER 50
FT NON_TER 50
SQ SEQUENCE 50 AA; 5448 MW; BBFC58B2511FC2E7 CRC64;

Query Match 75.6%; Score 31; DB 14; Length 50;
Best Local Similarity 71.4%; Pred. No. 22;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 MYPTYLK 8
DB 31 VWPTYLK 37

RESULT 14
OQX583 PRELIMINARY; PRT; 53 AA.
ID OQX583
AC OQX583;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE TUBBY PROTEIN (FRAGMENT).
GN TUB.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99160468; PubMed=10051323;
RA Caetano A.R., Pomp D., Murray J.D., Bowling A.T.;
RT "Comparative mapping of 18 equine type I genes assigned by somatic
RT cell hybrid analysis."
Mamm. Genome 10:271-276(1999).
EMBL: AF097580; AAD25983.1; -
HSSP: P50586; IC8Z.
DR InterPro: IPR000007; -
DR Pfam: PF01167; Tub; 1.
FT NON_TER 1
FT NON_TER 53
FT NON_TER 53
SQ SEQUENCE 53 AA; 6244 MW; 13A9ED8C69500D3E CRC64;

Query Match 75.6%; Score 31; DB 6; Length 53;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 MYPTY 6
DB 16 MYPTY 20

RESULT 15
OQ907 PRELIMINARY; PRT; 198 AA.
ID OQ907
AC OQ907;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)

DE GP064R.
GN S064R.
OS Shope fibroma virus (strain Kasza) (SFV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Leporipoxvirus.
OX NCBI_TaxID=10272;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-KASZA;
RX MEDLINE=84165064; PubMed=6323741;
RA Delange A.M., Macaulay C., Block W., Mueller T., McFadden G.;
RT "Tumorigenic poxviruses: construction of the composite physical map of
RT the Shope fibroma virus genome."
RL J. Virol. 50:408-416(1984).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-KASZA;
RX MEDLINE=92074222; PubMed=1660196;
RA Strayer D.S., Jerng H.H., O'Connor K.;
RT "Sequence and analysis of a portion of the genomes of Shope fibroma
RT virus and malignant rabbit fibroma virus that is important for viral
RT replication in lymphocytes."
RL Virology 185:585-595(1991).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-KASZA;
RX MEDLINE=20032074; PubMed=10562495;
RA Willer D.O., McFadden G., Evans D.H.;
RT "The complete genome sequence of shope (Rabbit) fibroma virus."
RL Virology 264:319-343(1999).
DR EMBL: AF170722; AAF17946.1; -
SQ SEQUENCE 198 AA; 23282 MW; 9D181052C1ED42C5 CRC64;

Query Match 75.6%; Score 31; DB 14; Length 198;
Best Local Similarity 83.3%; Pred. No. 96;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 MYPTYL 7
DB 110 MYPTFL 115

Search completed: September 4, 2001, 15:53:35
Job time: 288 sec

Tue Sep 4 15:56:19 2001

us-09-630-345-2.rspt

Page 6

0

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2001, 15:50:25 ; Search time 77.49 Seconds
(without alignments)
6.259 Million cell updates/sec

Title: US-09-630-345-3
Perfect score: 41
Sequence: 1 XMYPTYLK 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_0601.*
1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	97.6	8	19 AAW50937	Vasoactive intesti
2	40	97.6	8	21 AAB08298	Vasoactive intesti
3	40	97.6	8	21 AAB08306	Amino acid sequenc
4	35	85.4	8	21 AAB08304	Amino acid sequenc
5	35	85.4	8	21 AAB08305	Amino acid sequenc
6	35	85.4	450	21 AAB63104	Human secreted pro
7	32	78.0	807	21 AAY57288	Human GPCR protein
8	32	78.0	986	21 AAY40440	Human brain-derive
9	31	75.6	8	20 AAW75457	Mammalian tub prot
10	31	75.6	12	20 AAW97073	Peptidomimetic cap
11	31	75.6	58	20 AAY12222	Human 5' EST secre

12	31	75.6	173	22	AAW74618	Human 72 kDa gelat
13	31	75.6	173	22	AAW68682	Matrilysin catalyt
14	31	75.6	261	19	AAW76253	Human matrilysin-1
15	31	75.6	264	19	AAW76254	Human matrilysin p
16	31	75.6	271	17	AAW95211	Human recombinant
17	31	75.6	297	21	AAW43585	Human cancer assoc
18	31	75.6	325	14	AAW42750	ced-4 gene II prod
19	31	75.6	459	18	AAW36485	Mouse tub Form I
20	31	75.6	459	21	AAW26400	Mouse tub Form I p
21	31	75.6	460	18	AAW36488	Human tub Form 6
22	31	75.6	460	21	AAW26903	Human tub Form 6 p
23	31	75.6	462	22	AAW49983	Murine macrophage
24	31	75.6	469	21	AAW10655	BPV1 L1 fusion pro
25	31	75.6	469	21	AAW10660	BPV2 L1 fusion pro
26	31	75.6	483	20	AAW93163	Human MMP-20 prote
27	31	75.6	495	6	AAW51099	Sequence of bovine
28	31	75.6	495	20	AAW8482	Bovine papillomavi
29	31	75.6	505	18	AAW36486	Mouse tub Form II
30	31	75.6	505	18	AAW10728	Mouse tub gene pro
31	31	75.6	505	19	AAW54367	Mouse tub polypt
32	31	75.6	505	20	AAW75450	Mouse wild type tu
33	31	75.6	505	21	AAW26901	Mouse tub Form II
34	31	75.6	506	18	AAW36497	Human tub Form 4
35	31	75.6	506	18	AAW10729	Human tub gene pro
36	31	75.6	506	19	AAW54368	Human tub polypt
37	31	75.6	506	20	AAW75451	Human wild type tu
38	31	75.6	506	21	AAW26910	Human tub form 4 5
39	31	75.6	512	18	AAW36496	Human tub Form 3
40	31	75.6	512	21	AAW26912	Human tub form 3 5
41	31	75.6	518	18	AAW36495	Human tub Form 2
42	31	75.6	518	21	AAW26909	Human tub form 2 5
43	31	75.6	561	18	AAW36489	Human tub Form 1
44	31	75.6	561	21	AAW26904	Human tub Form 1 p
45	30	73.2	73	14	AAW40041	Fillistata peptide

ALIGNMENTS

RESULT 1
AAW50937
ID AAW50937 standard; peptide; 8 AA.
XX
AC AAW50937;
XX
DT 31-JUL-1998 (first entry)
XX
DE Vasoactive intestinal peptide receptor binding inhibitor (VIP2).
XX
KW Vasoactive intestinal peptide; VIP; antagonist; somatostatin; bombesin;
KW Substance P; cancer; inhibition.
XX
OS Synthetic.
XX
PN EP835662-A2.
XX
PD 15-APR-1998.
XX
PF 11-DEC-1996; 96EP-0309012.
XX
PR 08-OCT-1996; 96US-0727679.
PR 16-AUG-1996; 96IN-0001822.
XX
(NATIM-) NAT INST IMMUNOLOGY.
XX
Jaggi M, Mukherjee R;
XX
WPI; 1998-208959/19.
XX
Composition containing analogues of vasoactive intestinal peptide,
PT somatostatin - bombesin and substance P, for treatment of tumours
PT and for inhibiting over-expression of these peptide(s)

PS Claim 1; Page 4; 49pp; English.

XX The invention relates to a new composition which comprises: (i) the
CC somatostatin analogue SOM2 AGCKNFDKPTSDC (3-14 disulphide bridge),
CC and (ii) at least 4 of the peptides: antagonist of vasoactive
CC intestinal peptide (VIP1); VIP receptor-binding inhibitor (VIP2); VIP
CC receptor antagonist (VIP3); somatostatin analogue (SOM1); bombesin
CC antagonist (BOM1) and substance P antagonist (SP1). Also claimed are
CC more general compositions containing peptide analogues of somatostatin,
CC VIP, bombesin and substance P. The compositions are used in human or
CC veterinary medicine: (a) to kill (or inhibit multiplication of) tumour
CC or cancer cells, particularly for treatment of leukaemia, lymphoma,
CC adenocarcinoma of stomach, pancreas or prostate, or cancer of lung,
CC breast, kidney or particularly rectum and colon, and (b) to prevent,
CC inhibit or modulate over-expression of, e.g. VIP. A wide range of cancer
CC cells express receptors for VIP, somatostatin, bombesin and/or substance
CC P. The present sequence represents VIP receptor-binding inhibitor
CC (VIP2).

XX Sequence 8 AA;

Query Match 97.6%; Score 40; DB 19; Length 8;
Best Local Similarity 100.0%; Pred. No. 3.4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 MYPTYLK 8
Db 111111
2 myptylk 8

RESULT 2

AA08298
ID AAB08298 standard; peptide; 8 AA.

XX AAB08298;

DT 04-DEC-2000 (first entry)

XX Vasoactive intestinal peptide (VIP) analogue VIP2.

XX Vasoactive intestinal peptide; VIP; analogue; somatostatin; SOM1; SOM2;
KW VIP1; VIP2; VIP3; BOM1; bombesin; SP1; substance P; MuJ-7; tumour growth;
KW tumour angiogenesis; metastasis; cancer; angiogenesis; adenocarcinoma;
KW leukaemia; lymphoma.

XX Synthetic.

XX WO200047221-A1.

XX 17-AUG-2000.

XX 11-FEB-2000; 2000WO-US03559.

XX 11-FEB-1999; 99US-0248381.

XX (NAIN-) NAT INST IMMUNOLOGY.

XX (DABU-) DABUR RES FOUND.

XX (CORD/) CORD J I.

XX Mukherjee R, Jaggi M, Prasad S, Burman AC, Rajendran P, Mathur A;
PI Singh AT;

XX WPI; 2000-549083/50.

XX Novel therapeutically active composition comprising at least 5
PT peptides, useful for treating angiogenesis especially as a result of
PT adenocarcinomas.

XX Disclosure; Page 8; 42pp; English.

XX The present sequence represents an analogue of vasoactive intestinal
CC peptide (VIP). The specification describes therapeutically active

CC compositions comprising at least one analogue of somatostatin (chosen
CC from SOM1 and SOM2), and at least four analogues chosen from VIP1 (a
CC VIP antagonist), VIP2 (a VIP receptor binding inhibitor), VIP3 (a VIP
CC receptor antagonist), BOM1 (a bombesin antagonist), and SP1 (a substance
CC P antagonist). The combination of these 7 analogues is known as MuJ-7.
CC MuJ-7 is used as an anticancer drug to restrict tumour growth and spread
CC by inhibiting tumour angiogenesis. MuJ-7 in addition, inhibits
CC metastasis through its antiangiogenic activity in all cancers. The
CC peptides are useful for the treatment and prevention of angiogenesis,
CC especially as a result of adenocarcinomas of the colon, breast, lung,
CC prostate, kidney, leukemias or lymphomas.

XX Sequence 8 AA;

Query Match 97.6%; Score 40; DB 21; Length 8;
Best Local Similarity 100.0%; Pred. No. 3.4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 MYPTYLK 8
Db 111111
2 myptylk 8

RESULT 3

AA08306
ID AAB08306 standard; peptide; 8 AA.

XX AAB08306;

DT 04-DEC-2000 (first entry)

XX Amino acid sequence of an antiangiogenic peptide.

XX Vasoactive intestinal peptide; VIP; analogue; somatostatin; SOM1; SOM2;
KW VIP1; VIP2; VIP3; BOM1; bombesin; SP1; substance P; MuJ-7; tumour growth;
KW tumour angiogenesis; metastasis; cancer; angiogenesis; adenocarcinoma;
KW leukaemia; lymphoma.

XX Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 1 /note= "optionally D-form residue"

FT Misc-difference 3 /note= "optionally D-form residue"

FT Misc-difference 6 /note= "optionally D-form residue"

FT Misc-difference 7 /note= "optionally D-form residue"

FT /note= "optionally D-form residue"

XX WO200047221-A1.

XX 17-AUG-2000.

XX 11-FEB-2000; 2000WO-US03559.

XX 11-FEB-1999; 99US-0248381.

XX (NAIN-) NAT INST IMMUNOLOGY.

XX (DABU-) DABUR RES FOUND.

XX (CORD/) CORD J I.

XX Mukherjee R, Jaggi M, Prasad S, Burman AC, Rajendran P, Mathur A;
PI Singh AT;

XX WPI; 2000-549083/50.

XX Novel therapeutically active composition comprising at least 5
PT peptides, useful for treating angiogenesis especially as a result of
PT adenocarcinomas.

XX Claim 11; Page 31; 42pp; English.

XX AAB08304-15 represent peptides which have an antiangiogenic effect. The
 CC specification describes therapeutically active compositions comprising
 CC at least one analogue of somatostatin (chosen from SOM1 and SOM2), and
 CC at least four analogues chosen from vasoactive intestinal peptide (VIP)
 CC 1 (a VIP antagonist), VIP2 (a VIP receptor binding inhibitor), VIP3 (a
 CC VIP receptor antagonist), BOM1 (a bombesin antagonist), and SP1 (a
 CC substance P antagonist). The combination of these 7 analogues is known as
 CC MuJ-7. MuJ-7 is used as an anticancer drug to restrict tumour growth and
 CC spread by inhibiting tumour angiogenesis. MuJ-7, in addition, inhibits
 CC metastasis through its antiangiogenic activity in all cancers. The
 CC peptides are useful for the treatment and prevention of angiogenesis,
 CC especially as a result of adenocarcinomas of the colon, breast, lung,
 CC prostate, kidney, leukemias or lymphomas.
 XX SQ Sequence 8 AA;

Query Match 97.6%; Score 40; DB 21; Length 8;
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
 |||||
 Db 2 myptylk 8

RESULT 4
 AAB08304
 ID AAB08304 standard; peptide; 8 AA.
 XX AC AAB08304;
 XX DT 04-DEC-2000 (first entry)
 XX DE Amino acid sequence of an antiangiogenic peptide.
 XX KW Vasoactive intestinal peptide; VIP; analogue; somatostatin; SOM1; SOM2;
 KW VIP1; VIP2; VIP3; BOM1; bombesin; SP1; substance P; MuJ-7; tumour growth;
 KW tumour angiogenesis; metastasis; cancer; angiogenesis; adenocarcinoma;
 KW leukaemia; lymphoma.
 XX OS Synthetic.
 XX FH Key Location/Qualifiers
 FT Modified-site 1 /label= Aib
 FT /note= "alpha-aminoisobutyric acid"
 FT Modified-site 7 /label= Aib, Deg
 FT /note= "alpha-aminoisobutyric acid or alpha, alpha
 diethyl glycine"
 FT WO200047221-A1.
 XX PN 17-AUG-2000.
 XX PD 11-FEB-2000; 2000WO-US03559.
 XX PF 11-FEB-1999; 99US-0248381.
 XX PR (NAIM-) NAT INST IMMUNOLOGY.
 XX PA (DABU-) DABUR RES FOUND.
 XX PA (CORD/) CORD J I.
 XX PI Mukherjee R, Jaggi M, Prasad S, Burman AC, Rajendran P, Mathur A;
 PI Singh Ar;
 XX WPI; 2000-549083/50.
 XX DR Novel therapeutically active composition comprising at least 5
 XX peptides, useful for treating angiogenesis especially as a result of
 PT adenocarcinomas -

XX Claim 18; Page 36; 42pp; English.
 XX AAB08304-15 represent peptides which have an antiangiogenic effect. The
 CC specification describes therapeutically active compositions comprising
 CC at least one analogue of somatostatin (chosen from SOM1 and SOM2), and
 CC at least four analogues chosen from vasoactive intestinal peptide (VIP)
 CC 1 (a VIP antagonist), VIP2 (a VIP receptor binding inhibitor), VIP3 (a
 CC VIP receptor antagonist), BOM1 (a bombesin antagonist), and SP1 (a
 CC substance P antagonist). The combination of these 7 analogues is known as
 CC MuJ-7. MuJ-7 is used as an anticancer drug to restrict tumour growth and
 CC spread by inhibiting tumour angiogenesis. MuJ-7, in addition, inhibits
 CC metastasis through its antiangiogenic activity in all cancers. The
 CC peptides are useful for the treatment and prevention of angiogenesis,
 CC especially as a result of adenocarcinomas of the colon, breast, lung,
 CC prostate, kidney, leukemias or lymphomas.
 XX SQ Sequence 8 AA;

Query Match 85.4%; Score 35; DB 21; Length 8;
 Best Local Similarity 85.7%; Pred. No. 3.4e+05;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
 |||||
 Db 2 myptyxx 8

RESULT 5
 AAB08305
 ID AAB08305 standard; peptide; 8 AA.
 XX AC AAB08305;
 XX DT 04-DEC-2000 (first entry)
 XX DE Amino acid sequence of an antiangiogenic peptide.
 XX KW Vasoactive intestinal peptide; VIP; analogue; somatostatin; SOM1; SOM2;
 KW VIP1; VIP2; VIP3; BOM1; bombesin; SP1; substance P; MuJ-7; tumour growth;
 KW tumour angiogenesis; metastasis; cancer; angiogenesis; adenocarcinoma;
 KW leukaemia; lymphoma.
 XX OS Synthetic.
 XX FH Key Location/Qualifiers
 FT Misc-difference 1 /note= "D-form residue"
 FT Modified-site 7 /label= Aib, Deg
 FT /note= "alpha-aminoisobutyric acid or alpha, alpha
 diethyl glycine"
 FT WO200047221-A1.
 XX PN 17-AUG-2000.
 XX PD 11-FEB-2000; 2000WO-US03559.
 XX PF 11-FEB-1999; 99US-0248381.
 XX PR (NAIM-) NAT INST IMMUNOLOGY.
 XX PA (DABU-) DABUR RES FOUND.
 XX PA (CORD/) CORD J I.
 XX PI Mukherjee R, Jaggi M, Prasad S, Burman AC, Rajendran P, Mathur A;
 PI Singh Ar;
 XX WPI; 2000-549083/50.
 XX DR Novel therapeutically active composition comprising at least 5
 XX peptides, useful for treating angiogenesis especially as a result of
 PT adenocarcinomas -

adenocarcinomas -

Claim 11; Page 31; 42pp; English.

AA08304-15 represent peptides which have an antiangiogenic effect. The specification describes therapeutically active compositions comprising at least one analogue of somatostatin (chosen from SOM1 and SOM2), and at least four analogues chosen from vasoactive intestinal peptide (VIP) 1 (a VIP antagonist), VIP2 (a VIP receptor binding inhibitor), VIP3 (a VIP receptor antagonist), BOM1 (a bombesin antagonist), and SP1 (a substance P antagonist). The combination of these 7 analogues is known as MuJ-7. MuJ-7 is used as an anticancer drug to restrict tumour growth and spread by inhibiting tumour angiogenesis. MuJ-7, in addition, inhibits metastasis through its antiangiogenic activity in all cancers. The peptides are useful for the treatment and prevention of angiogenesis, especially as a result of adenocarcinomas of the colon, breast, lung, prostate, kidney, leukemias or lymphomas.

Sequence 8 AA;

Query Match 85.4%; Score 35; DB 21; Length 8;
Best Local Similarity 85.7%; Pred. No. 3.4e+05;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
| | | | |
Db 2 myptyxk 8

RESULT 6
AA063104
ID AAB63104 standard; Protein; 450 AA.

AC AAB63104;

DT 26-MAR-2001 (first entry)

DE Human secreted protein sequence encoded by gene 18 SEQ ID NO:114.

XX Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;
KW antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;
KW cerebroprotective; neurotropic; neuroprotective; antibacterial; virucide;
KW fungicide; ophthalmological; vulnery; gene therapy; neoplasm;
KW autoimmune disease; rheumatoid arthritis; hyperproliferative disorder;
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW cerebral ischaemia; angiogenesis; nervous system disorder; infection;
KW Alzheimer's disease; ocular disorder; corneal infection; wound healing;
skin aging; food additive; preservative.

OS Homo sapiens.

PN WO200061748-A1.

XX 19-OCT-2000.

XX 06-APR-2000; 2000WO-US08982.

XX 09-APR-1999; 99US-0128696.

XX 14-JAN-2000; 2000US-0176069.

PA (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM, Komatsoulis G;

XX WPI; 2000-638566/61.

XX New nucleic acid molecules encoding 48 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions and
PT used as food additives or preservatives.

XX Disclosure; Page 453-455; 480pp; English.

AA022316 to AAF22363 encode the human secreted proteins given in AAB63049 to AAB63096. AAB63097 to AAB63132 represent more human secreted proteins and polypeptides homologous to them. Human secreted proteins have activities based on the tissues and cells the genes are expressed in. Examples of activities include: immunosuppressive; antiarthritic; antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective; neurotropic; neuroprotective; antibacterial; virucide; fungicide; ophthalmological; and vulnery. The polynucleotides and proteins can be used to prevent, treat or ameliorate a medical condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used in diagnosing a pathological condition or susceptibility to a pathological condition. Disorders which are diagnosed or treated include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g. Alzheimer's disease, infections caused by bacteria, viruses and fungi and ocular disorders e.g. corneal infection. The polypeptides can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues and in chemotaxis. The polypeptides can also be used as a food additive or preservative to increase or decrease storage capabilities, fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors and other nutritional components. AAF22307 to AAF22315 and AAB63048 represent sequences used in the exemplification of the present invention.

Sequence 450 AA;

Query Match 85.4%; Score 35; DB 21; Length 450;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYL 7
| | | | |
Db 106 myptyl 111

RESULT 7
AA057288
ID AAY57288 standard; Protein; 807 AA.

AC AAY57288;

DT 05-JUN-2000 (first entry)

DE Human GPCR protein (HGPRP) sequence (clone ID 3036563).

XX Human; G protein coupled protein receptor; HGPRP; cell proliferation;
KW neurological; immune disorder; cytostatic; anti-arteriosclerotic;
KW anti-atherosclerotic; hepatotropic; antinflammatory; virucide; leukemia;
KW immunomodulatory; anemia; asthma; gastrointestinal; anti-epileptic;
KW anti-Alzheimer's; anti-Parkinsonian; gene therapy.

OS Homo sapiens.

PN WO200015793-A2.

XX 23-MAR-2000.

XX 17-SEP-1999; 99WO-US20958.

XX 17-SEP-1998; 98US-0156513.

XX (INCY-) INCYTE PHARM INC.

XX Bandman O, Lal P, Tang YT, Corley NC, Guegler KJ, Gorgone GA;
PI Baughn MR;

XX WPI: 2000-271432/23.

XX DR N-PSDB; AA290526.

PT Human G protein coupled protein receptor peptides useful for the
PT prevention, diagnosis and treatment of cell proliferative, neurological
XX and immune disorders -

XX Claim 1; Page 65-67; 71pp; English.

XX The invention provides human G protein coupled protein receptor (HGPRP)
XX polypeptides and polynucleotides encoding them. The polypeptides can be
CC produced by standard recombinant methodology. The polynucleotides and
CC polypeptides may be used in the prevention, treatment and diagnosis of
CC diseases associated with their inappropriate expression. Diseases that
CC can be treated are cell proliferative disorders (e.g. arteriosclerosis,
CC atherosclerosis and hepatitis), cancers (e.g. leukemia, melanomas and
CC adenocarcinoma), immune disorders (e.g. anemia, asthma and Crohn's
CC disease) and neurological disorders (e.g. epilepsy, Alzheimer's disease
CC and Parkinson's disease). The anti-HGPRP antibodies may also be used as
CC diagnostic agents for detecting the presence of HGPRP polypeptides in
CC samples (e.g. by enzyme linked immunosorbant assay (ELISA)). Sequences
CC AAY57283-288 represent the HGPRP polypeptides.

Sequence 807 AA;

Query Match 78.0%; Score 32; DB 21; Length 807;
Best Local Similarity 85.7%; Pred. No. 2.9e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
I | | | | |
Db 205 mlptylk 211

RESULT 8

AAAY40440
ID AAY40440 standard; Protein; 986 AA.

XX AC AAY40440;

XX DT 26-MAY-2000 (first entry)

XX DE Human brain-derived G-protein coupled receptor protein.

XX KW G-protein coupled receptor protein; human; brain; gene therapy;
XX KW genetic disease; screening assay.

XX OS Homo sapiens.

XX WO200008053-A1.

XX 17-FEB-2000.

XX PF 05-AUG-1999; 99WO-JP04233.

XX PR 07-AUG-1998; 98JP-0225059.

XX PA (TAKE) TAKEDA CHEM IND LTD.
XX PA (KAZU-) KAZUSA DNA RES INST.

XX PI Ohara O, Nagase T, Nomura N, Mogi S, Yamamoto K, Kurokawa T;

XX WPI; 2000-195555/17.

XX N-PSDB; AAZ87930.

XX Human brain-derived G-protein coupled receptor protein and encoding DNA
PT used in gene therapy, is also useful for e.g. identifying ligands and
PT raising antibodies and antisera -

XX Claim 1; Fig 1; 80pp; Japanese.

XX This represents a human brain-derived G-protein coupled receptor
CC protein. The polypeptide can be expressed by standard recombinant
CC methodology. The novel G-protein coupled receptor protein can be used
CC for identifying ligands; raising antibodies and antisera; developing

CC receptor-binding assay system; and screening for drug candidates. The
CC encoding polynucleotide can be used in; probing for diagnostic genes;
CC constructing PCR primers; making transgenic animals; and in gene therapy.
XX
SQ Sequence 986 AA;

Query Match 78.0%; Score 32; DB 21; Length 986;
Best Local Similarity 85.7%; Pred. No. 3.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
I | | | | |
Db 384 mlptylk 390

RESULT 9

AAAW75457

ID AAW75457 standard; peptide; 8 AA.

XX AC AAW75457;

XX DT 27-APR-1999 (first entry)

XX DE Mammalian tub protein tyrosine phosphorylation site.

XX KW Human; wild type; tubby; identification; SH2 domain; mammal; obesity;
XX KW body weight disorder; cachexia; anorexia.

XX OS Homo sapiens.

XX OS Mus sp.

XX PN US5861239-A.

XX PD 19-JAN-1999.

XX PF 02-SEP-1997; 97US-0922267.

XX PR 02-SEP-1997; 97US-0922267.

XX PR 12-APR-1996; 96US-0631200.

XX PR 28-MAR-1997; 97US-0829553.

XX PA (MILL-) MILLENNIUM PHARM INC.

XX PI Kapeller R, Kleyn PW, Moore KJ;

XX WPI; 1999-130383/11.

XX PT Identifying compounds which modulate tub protein activity - by
PT detecting compounds which alter the interaction of tub protein with
PT a SH2 containing peptide, used to develop agents for treating e.g.
PT obesity, cachexia or anorexia

XX PS Disclosure; Column 13; 95pp; English.

XX This sequence represents a tyrosine phosphorylation site/SH2 docking
CC domain found in the amino acid sequence of the mouse and human "tub"
CC proteins (AAW75450 and AAW75451 respectively). The invention relates to
CC a method for identifying compounds that modulate tub protein activity,
CC especially its interaction with proteins containing an SH2 domain. The
CC method can be used for identifying compounds which modulate tub protein
CC activity for use in the treatment of mammalian body weight disorders
CC including obesity, cachexia and anorexia.

XX SQ Sequence 8 AA;

Query Match 75.6%; Score 31; DB 20; Length 8;
Best Local Similarity 100.0%; Pred. No. 3.4e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
I | | | | |

Db 2 mypty 6

RESULT 10

AAW97073
ID AAW97073 standard; peptide; 12 AA.

AC AAW97073;

XX 29-APR-1999 (first entry)

XX Peptidomimetic capable of inhibiting CD28 and/or CTLA-4 interactions.

XX Peptidomimetic; CD28; CTLA-4; CD80; CD86; B7-1; B7-2; diabetes;
XX immune system disease; autoimmune disease; psoriasis; multiple sclerosis;
XX lupus erythematosus; rheumatoid arthritis; transplant rejection; cancer.

XX Synthetic.

XX WO9856401-A1.

XX 17-DEC-1998.

XX 11-JUN-1998; 98WO-US12312.

XX 12-JUN-1997; 97US-0049470.

XX (ISTF) ARS APPLIED RES SYSTEMS HOLDING NV.

XX Blechner S, El Tayar N, Jameson B, Tepper M;

XX WPI; 1999-080854/07.

XX New peptidomimetic compounds - useful in the prophylaxis, diagnosis
XX and treatment of pathologies and disorders, which are improved by
XX inhibition of CD28 and/or CTLA-4 interaction with CD80 (B7-1) and
XX CD86 (B7-2)

XX Disclosure; Page 9; 62pp; English.

XX AAW97053-82 represent peptidomimetic compounds that inhibit CD28 and/or
XX CTLA-4 interactions with CD80 and CD86. The peptides are used in
XX pharmaceutical compositions for the prophylaxis, diagnosis and treatment
XX of pathologies and disorders, which are improved by inhibition of CD28
XX and/or CTLA-4 interaction with CD80 (B7-1) and CD86 (B7-2). Disorders
XX include immune system diseases (e.g. Autoimmune diseases such as
XX psoriasis, multiple sclerosis, lupus erythematosus, diabetes, rheumatoid
XX arthritis, and therapy against solid organ/cellular transplant
XX rejection) and cancer.

XX Sequence 12 AA;

XX Query Match 75.68; Score 31; DB 20; Length 12;

XX Best Local Similarity 71.48; Pred. No. 5.2;

XX Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 MYPTYLK 8

DB 2 myptylr 8

XX 18-JUN-1999 (first entry)

XX Human 5' EST secreted protein SEQ ID NO: 535.

XX AAY12222

XX ID AAY12222 standard; Protein; 58 AA.

XX AC AAY12222;

XX 18-JUN-1999 (first entry)

XX Human 5' EST secreted protein SEQ ID NO: 535.

XX AAY12222

XX ID AAY12222 standard; Protein; 58 AA.

XX AC AAY12222;

XX 18-JUN-1999 (first entry)

XX Human 5' EST secreted protein SEQ ID NO: 535.

XX AAY12222

XX ID AAY12222 standard; Protein; 58 AA.

XX AC AAY12222;

XX 18-JUN-1999 (first entry)

XX Human 5' EST secreted protein SEQ ID NO: 535.

XX AAY12222

XX ID AAY12222 standard; Protein; 58 AA.

XX AC AAY12222;

XX 18-JUN-1999 (first entry)

XX Human 5' EST secreted protein SEQ ID NO: 535.

XX AAY12222

XX ID AAY12222 standard; Protein; 58 AA.

XX AC AAY12222;

XX 18-JUN-1999 (first entry)

XX Human 5' EST secreted protein SEQ ID NO: 535.

XX AAY12222

XX ID AAY12222 standard; Protein; 58 AA.

XX AC AAY12222;

XX 18-JUN-1999 (first entry)

XX Human 5' EST secreted protein SEQ ID NO: 535.

XX AAY12222

XX ID AAY12222 standard; Protein; 58 AA.

XX AC AAY12222;

XX 18-JUN-1999 (first entry)

XX Human 5' EST secreted protein SEQ ID NO: 535.

XX AAY12222

KW Human; secreted protein; EST; expressed sequence tag; diagnosis;
 KW forensic; gene therapy; chromosome mapping; signal peptide;
 KW upstream regulatory sequence; cytokine activity; cell proliferation;
 KW differentiation; haematopoiesis regulation; tissue growth regulation;
 KW reproductively hormone regulation; chemotactic; chemokinetic; haemostatic;
 KW thrombolytic; anti-inflammatory; tumour inhibition.

XX Homo sapiens.

XX WO9806554-A2.

XX 11-FEB-1999.

XX 31-JUL-1998; 98WO-IB01238.

XX 01-AUG-1997; 97US-0905134.

XX (GEST) GENSET.

XX Duclert A, Dumas Milne Edwards J, Lacroix B;

XX WPI; 1999-153784/13.

XX N-PSDB; AAX41055.

XX New nucleic acids encoding human secreted proteins - obtained from
 XX cDNA libraries prepared from kidney, fetal kidney, dystrophic
 XX muscle, muscle and heart tissue

XX Claim 34; Page 598; 622pp; English.

XX AAX40826 to AAX41093 represent 5' expressed sequence tags (ESTs) for
 XX human secreted proteins, and encode the proteins given in AAY01602 and
 XX AAY11994 to AAY12260, respectively. The proteins given represent the
 XX signal peptide and an N-terminal fragment of a secreted protein. The
 XX nucleic acid sequences can be used for producing secreted protein. The
 XX products may also be used to develop products for diagnosis and
 XX therapy. The proteins obtained may have cytokine activity, cell
 XX proliferation/differentiation activity, haematopoiesis regulating
 XX activity, tissue growth regulating activity, reproductively hormone
 XX regulating activity, chemotactic/chemokinetic activity, haemostatic and
 XX thrombolytic activity, receptor/ligand activity, anti-inflammatory
 XX activity, tumour activity or other activities. The products
 XX can be used in forensic, gene therapy and chromosome mapping procedures.
 XX The sequences can also be used for obtaining corresponding promoter
 XX sequences. The nucleic acids encoding the signal peptide can be used
 XX for directing extracellular secretion of a polypeptide or the insertion
 XX of a polypeptide into a membrane, or importing a polypeptide into
 XX a cell.

XX Sequence 58 AA;

XX Query Match 75.68; Score 31; DB 20; Length 58;

XX Best Local Similarity 83.38; Pred. No. 28;

XX Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYL 7

DB 1 mypsyl 6

XX RESULT 12

XX AAB74618

XX ID AAB74618 standard; protein; 173 AA.

XX AC AAB74618;

XX 22-MAY-2001 (first entry)

XX Human 72 kDa gelatinase catalytic domain protein sequence SEQ ID NO:8.

XX Human; 92 kDa gelatinase; 72 kDa gelatinase; matrilysin; gelatinase;
 XX catalytic domain; 92 CD; 72 CD; matrilysin CD; elastin; vulnery;

XX AAB74618

XX ID AAB74618 standard; protein; 173 AA.

XX AC AAB74618;

XX 22-MAY-2001 (first entry)

XX Human 72 kDa gelatinase catalytic domain protein sequence SEQ ID NO:8.

XX Human; 92 kDa gelatinase; 72 kDa gelatinase; matrilysin; gelatinase;
 XX catalytic domain; 92 CD; 72 CD; matrilysin CD; elastin; vulnery;

XX AAB74618

XX ID AAB74618 standard; protein; 173 AA.

XX AC AAB74618;

XX 22-MAY-2001 (first entry)

XX Human 72 kDa gelatinase catalytic domain protein sequence SEQ ID NO:8.

XX Human; 92 kDa gelatinase; 72 kDa gelatinase; matrilysin; gelatinase;
 XX catalytic domain; 92 CD; 72 CD; matrilysin CD; elastin; vulnery;

XX AAB74618

XX ID AAB74618 standard; protein; 173 AA.

XX AC AAB74618;

XX 22-MAY-2001 (first entry)

XX Human 72 kDa gelatinase catalytic domain protein sequence SEQ ID NO:8.

XX Human; 92 kDa gelatinase; 72 kDa gelatinase; matrilysin; gelatinase;
 XX catalytic domain; 92 CD; 72 CD; matrilysin CD; elastin; vulnery;

XX AAB74618

XX ID AAB74618 standard; protein; 173 AA.

XX AC AAB74618;

XX 22-MAY-2001 (first entry)

XX Human 72 kDa gelatinase catalytic domain protein sequence SEQ ID NO:8.

XX Human; 92 kDa gelatinase; 72 kDa gelatinase; matrilysin; gelatinase;
 XX catalytic domain; 92 CD; 72 CD; matrilysin CD; elastin; vulnery;

XX AAB74618

XX ID AAB74618 standard; protein; 173 AA.

XX AC AAB74618;

XX 22-MAY-2001 (first entry)

XX Human 72 kDa gelatinase catalytic domain protein sequence SEQ ID NO:8.

XX Human; 92 kDa gelatinase; 72 kDa gelatinase; matrilysin; gelatinase;
 XX catalytic domain; 92 CD; 72 CD; matrilysin CD; elastin; vulnery;

XX AAB74618

XX ID AAB74618 standard; protein; 173 AA.

XX AC AAB74618;

XX 22-MAY-2001 (first entry)

XX Human 72 kDa gelatinase catalytic domain protein sequence SEQ ID NO:8.

KW excess connective tissue removal; dermatological; keloid; scleroderma;
 KW post-operative fibrosis; intervertebral disc injection; fibrotic disease;
 KW hypertrophic scar; wound debridement; post-surgical adhesion;
 XX idiopathic pulmonary fibrosis.

OS Homo sapiens.

XX US6194189-B1.

XX 27-FEB-2001.

XX 16-DEC-1994; 94US-0357820.

XX 16-DEC-1994; 94US-0357820.

XX (UNIW) UNIV WASHINGTON.

XX Senior RM;

XX WPI; 2001-243407/25.

PT New gelatinase truncated mutant useful for treating disorders requiring
 PT the removal of excess connective tissue, e.g. keloids, post-operative
 PT fibrosis, intervertebral disc injections, hypertrophic scars -

XX Example; Column 13-16; 11pp; English.

XX The present invention describes a cDNA sequence which encodes the
 CC truncated mutant of the 92 kDa gelatinase having an amino acid sequence
 CC as given in AAB74617 consisting of residues 106-216 fused to residues
 CC 391-443 of the parent molecule. The truncated mutant has dermatological
 CC and vulnary activities. The truncated mutant is useful for treating
 CC disorders requiring the removal of excess connective tissue,
 CC e.g., keloids, post-operative fibrosis, intervertebral disc injections,
 CC hypertrophic scars, wound debridement, post-surgical adhesions and
 CC various fibrotic diseases (including scleroderma, idiopathic pulmonary
 CC fibrosis). The truncated mutant is catalytically active compared to
 CC the full protein. Unlike the full protein, the truncated mutant is
 CC essentially inactive against insoluble elastin, and does not require
 CC activation to be enzymatically active. The present sequence represents
 CC a 72 kDa gelatinase catalytic domain (72 CD) containing protein, which
 CC is given in the exemplification of the present invention.

XX Sequence 173 AA;

Query Match 75.6%; Score 31; DB 22; Length 173;
 Best Local Similarity 100.0%; Pred. No. 89;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 MYPTY 6

DB 138 mypty 142

RESULT 13

AAB68682

ID AAB68682 standard; Protein; 173 AA.

XX AAB68682;

XX 02-MAY-2001 (first entry)

XX Matrilysin catalytic domain.

XX Gelatinase; excess connective tissue removal; matrilysin;
 KW matrix metalloproteinase.

XX Unidentified.

XX US6184021-B1.

XX 06-FEB-2001.

XX

PF 19-MAY-1995; 95US-0444628.

XX 16-DEC-1994; 94US-0357820.

XX (UNIW) UNIV WASHINGTON.

XX Senior RM;

XX WPI; 2001-202001/20.

XX New truncated mutant of 92 kDa gelatinase which is catalytically
 PT active, but is inactive against insoluble elastin, useful for treating
 PT disorders requiring the removal of excess connective tissues such as
 PT keloids -

XX Examples; Fig 2; 11pp; English.

XX The present invention relates to a truncated mutant (92 CD) of the 92
 CC kDa gelatinase (see AAB68681). The truncated protein is
 CC useful for treating disorders requiring the removal of excess connective
 CC tissue, e.g. keloids, post-operative fibrosis, intervertebral disc
 CC injections, hypertrophic scars, wound debridement, post-surgical
 CC adhesions and various fibrotic diseases (scleroderma, idiopathic
 CC pulmonary fibrosis). Gelatinase is a matrix metalloproteinase and is also
 CC known as gelatinase B and MMP-9. The truncated protein is catalytically
 CC active comparable to the full protein but unlike the full protein is
 CC essentially inactive against insoluble elastin. The present sequence is
 CC the catalytic domain of matrilysin, which was used in a sequence homology
 CC alignment with the 92 CD protein of the present invention. Matrilysin is
 CC also a matrix metalloproteinase.

XX Sequence 173 AA;

Query Match 75.6%; Score 31; DB 22; Length 173;

Best Local Similarity 100.0%; Pred. No. 89;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 MYPTY 6

DB 138 mypty 142

RESULT 14

AAW76253

ID AAW76253 standard; Protein; 261 AA.

XX AAW76253;

XX 02-DEC-1998 (first entry)

XX Human matrilysin-like protein.

XX Matrilysin; inhibitor; tumour necrosis factor; screening; MMP; infection;
 KW matrix metalloproteinase; antagonist; diagnosis; underexpression; cancer;
 KW degeneration; extracellular matrix; arthritis; cardiovascular disease;
 KW cachexia; multiple sclerosis.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..22

FT /label= signal

FT Protein 23..261

FT /label= matrilysin_like_protein

XX WO9831818-A2.

XX 23-JUL-1998.

XX 20-JAN-1998; 98WO-US00783.

XX

```

PR 01-AUG-1997; 97US-0054541.
PR 21-JAN-1997; 97US-0034205.
PR 13-JUN-1997; 97US-0049607.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Brewer L, Gentz R, Ni J, Rosen CA, Ruben SM;
XX WPI: 1998-414114/35.
XX DR N-PSDB; AAV61633.
XX
XX Isolated nucleic acid encoding human metallo-protease(s) - used for
XX diagnosis, treatment and prevention of, e.g. cancer, inflammation,
XX neurological disease and infections
XX Claim 20a; Fig 4; 81pp; English.
XX
XX This sequence represents a novel human matrilysin-like protein which is
XX an inhibitor of the members of the matrix metalloproteinase, MMP,
XX family. This protein can be used in assays to screen for agonists and
XX antagonists and the nucleic acid is used as a probe for genes in
XX situ hybridisation and detection of corresponding genes in human tissue,
XX and as sources of probes and primers for diagnosis. The protein and its
XX antigenic fragments are used to raise antibodies (Ab) (which can be used
XX for diagnosis in usual immunoassays or for in vivo imaging) and to screen
XX for (antagonists. Agonists can be used to treat cancer or other to screen
XX conditions associated with underexpression of TNF-alpha. Antagonists are
XX used to treat diseases associated with degeneration of the extracellular
XX matrix (cancer, arthritis, cardiovascular disease, cachexia and multiple
XX sclerosis).
XX Sequence 261 AA;

Query Match 75.6%; Score 31; DB 19; Length 261;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
DB 226 mypty 230

RESULT 15
AAW76254
ID AAW76254 standard; Protein; 264 AA.
XX
XX AC AAW76254;
XX
XX 02-DEC-1998 (first entry)
XX Human matrilysin protein.
XX
XX Matrilysin; inhibitor; tumour necrosis factor; screening; MMP; infection;
XX matrix metalloproteinase; antagonist; diagnosis; underexpression; cancer;
XX degeneration; extracellular matrix; arthritis; cardiovascular disease;
XX cachexia; multiple sclerosis.
XX
XX Homo sapiens.
XX
XX WO9831818-A2.
XX
XX 23-JUL-1998.
XX
XX 20-JAN-1998; 98WO-US00783.
XX
XX 01-AUG-1997; 97US-0054541.
XX 21-JAN-1997; 97US-0034205.
XX 13-JUN-1997; 97US-0049607.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Brewer L, Gentz R, Ni J, Rosen CA, Ruben SM;

```

```

XX
XX WPI: 1998-414114/35.
XX
XX Isolated nucleic acid encoding human metallo-protease(s) - used for
XX diagnosis, treatment and prevention of, e.g. cancer, inflammation,
XX neurological disease and infections
XX Disclosure; Fig 5; 81pp; English.
XX
XX This sequence represents the human matrilysin protein which is an
XX inhibitor of the members of the matrix metalloproteinase, MMP, family.
XX This protein is used in the identification of a novel human
XX matrilysin-like protein which can be used in assays to screen for
XX agonists and antagonists and the nucleic acid is used as a probe for gene
XX mapping, in situ hybridisation and detection of corresponding genes in
XX human tissue, and as sources of probes and primers for diagnosis. The
XX protein and its antigenic fragments are used to raise antibodies (Ab)
XX (which can be used for diagnosis in usual immunoassays or for in vivo
XX imaging) and to screen for (antagonists. Agonists can be used to treat
XX cancer or other conditions associated with underexpression of TNF-alpha.
XX Antagonists are used to treat diseases associated with degeneration of
XX the extracellular matrix (cancer, arthritis, cardiovascular disease,
XX cachexia and multiple sclerosis).
XX Sequence 264 AA;

Query Match 75.6%; Score 31; DB 19; Length 264;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
DB 232 mypty 236

```

Search completed: September 4, 2001, 15:50:25
Job time: 163 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2001, 15:51:12 ; Search time 40.98 Seconds
(without alignments)
4.020 Million cell updates/sec

Title: US-09-630-345-3
Perfect score: 41
Sequence: 1 XMYPTYLK 8

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A-COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B-COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A-COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B-COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCTUS-COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	97.6	8	1 US-07-620-410-3	Sequence 3, Appli
2	31	75.6	8	2 US-08-922-267A-65	Sequence 65, Appl
3	31	75.6	173	4 US-08-444-628-8	Sequence 8, Appli
4	31	75.6	173	4 US-08-357-820-8	Sequence 8, Appli
5	31	75.6	261	3 US-09-009-156-5	Sequence 5, Appli
6	31	75.6	264	3 US-09-009-156-6	Sequence 6, Appli
7	31	75.6	267	4 US-08-448-489-18	Sequence 18, Appl
8	31	75.6	271	3 US-08-896-062-2	Sequence 2, Appli
9	31	75.6	459	1 US-08-630-592-2	Sequence 2, Appli
10	31	75.6	459	1 US-08-714-991-2	Sequence 2, Appli
11	31	75.6	459	3 US-09-032-365A-2	Sequence 2, Appli
12	31	75.6	460	1 US-08-630-592-7	Sequence 7, Appli
13	31	75.6	460	1 US-08-714-991-7	Sequence 7, Appli
14	31	75.6	460	3 US-09-032-365A-8	Sequence 8, Appli
15	31	75.6	462	4 US-08-068-392-3	Sequence 3, Appli
16	31	75.6	462	4 US-08-396-988-3	Sequence 3, Appli
17	31	75.6	505	1 US-08-631-200-2	Sequence 2, Appli
18	31	75.6	505	1 US-08-630-592-4	Sequence 4, Appli
19	31	75.6	505	1 US-08-714-991-4	Sequence 4, Appli
20	31	75.6	505	2 US-08-829-553-2	Sequence 2, Appli
21	31	75.6	505	2 US-08-922-267A-2	Sequence 2, Appli
22	31	75.6	505	2 US-08-936-707A-2	Sequence 2, Appli
23	31	75.6	505	2 US-08-936-706A-2	Sequence 2, Appli
24	31	75.6	505	3 US-09-248-203-2	Sequence 2, Appli
25	31	75.6	505	3 US-09-032-365A-4	Sequence 4, Appli
26	31	75.6	505	4 US-08-812-824-3	Sequence 3, Appli
27	31	75.6	505	4 US-09-406-071-2	Sequence 2, Appli

28	31	75.6	506	1 US-08-631-200-8	Sequence 8, Appli
29	31	75.6	506	2 US-08-829-553-8	Sequence 8, Appli
30	31	75.6	506	2 US-08-922-267A-8	Sequence 8, Appli
31	31	75.6	506	2 US-08-936-707A-8	Sequence 8, Appli
32	31	75.6	506	2 US-08-936-706A-8	Sequence 8, Appli
33	31	75.6	506	3 US-09-248-203-8	Sequence 8, Appli
34	31	75.6	506	3 US-09-032-365A-62	Sequence 62, Appli
35	31	75.6	506	4 US-08-812-824-4	Sequence 4, Appli
36	31	75.6	506	4 US-09-406-071-8	Sequence 8, Appli
37	31	75.6	512	3 US-09-032-365A-60	Sequence 60, Appli
38	31	75.6	518	3 US-09-032-365A-58	Sequence 58, Appli
39	31	75.6	561	1 US-08-714-991-27	Sequence 27, Appli
40	31	75.6	561	3 US-09-032-365A-10	Sequence 10, Appli
41	30	73.2	73	1 US-08-379-538-7	Sequence 7, Appli
42	30	73.2	452	2 US-08-731-079A-2	Sequence 2, Appli
43	30	73.2	547	4 US-09-178-252-11	Sequence 11, Appli
44	30	73.2	725	1 US-08-448-170-4	Sequence 4, Appli
45	30	73.2	725	4 US-08-961-803-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1
US-07-620-410-3
; Sequence 3, Application US/07620410
; Patent No. 5217953
; GENERAL INFORMATION:
; APPLICANT: Gozes, Iillana
; APPLICANT: Brennenman, Douglas E.
; APPLICANT: Fridkin, Matti
; APPLICANT: Moody, Terry
; TITLE OF INVENTION: A NOVEL VASOACTIVE INTESTINAL PEPTIDE
; TITLE OF INVENTION: ANTAGONIST
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
; STREET: 1615 L. Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/620,410
; FILING DATE: 19901130
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Scott, Watson T.
; REGISTRATION NUMBER: 26,581
; REFERENCE/DOCKET NUMBER: WTS/5683/82679
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 248453 CUSH
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
US-07-620-410-3

Query Match 97.6%; Score 40; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
|||||||

Db 2 MYPYLK 8

```
RESULT 2
US-08-922-267A-65
; Sequence 65, Application US/08922267A
; Patent No. 5861239
; GENERAL INFORMATION:
; APPLICANT: Klevn, Patrick W.
; APPLICANT: Moore, Karen J.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; NUMBER OF SEQUENCES: 82
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/922.267A
; FILING DATE: 2-SEP-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/829,553
; FILING DATE: 28-MAR-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/631,200
; FILING DATE: 12-APR-1996
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-085
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-922-267A-65

Query Match 75.6%; Score 31; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
DB 2 MYPTY 6

RESULT 3
US-08-444-628-8
; Sequence 8, Application US/08444628
; Patent No. 6184021
; GENERAL INFORMATION:
; APPLICANT: Senior, Robert M.
; TITLE OF INVENTION: Catalytically-active Gelatinase Mutant
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scott J. Meyer, Monsanto/Searle, A3SG
; STREET: 800 N. Lindbergh Blvd.
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63167
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/357,820
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyer, Scott J.
; REGISTRATION NUMBER: 25,275
; REFERENCE/DOCKET NUMBER: WU-2855
; TELECOMMUNICATION INFORMATION:
```

```
ADDRESSEE: Scott J. Meyer, Monsanto/Searle, A3SG
STREET: 800 N. Lindbergh Blvd.
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63167
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/444,628
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/357,820
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Meyer, Scott J.
REGISTRATION NUMBER: 25,275
REFERENCE/DOCKET NUMBER: WU-2855
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314)694-3117
TELEFAX: (314)694-5435
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 173 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-444-628-8

Query Match 75.6%; Score 31; DB 4; Length 173;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
DB 138 MYPTY 142

RESULT 4
US-08-357-820-8
; Sequence 8, Application US/08357820
; Patent No. 6194189
; GENERAL INFORMATION:
; APPLICANT: Senior, Robert M.
; TITLE OF INVENTION: Catalytically-active Gelatinase Mutant
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scott J. Meyer, Monsanto/Searle, A3SG
; STREET: 800 N. Lindbergh Blvd.
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63167
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/357,820
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyer, Scott J.
; REGISTRATION NUMBER: 25,275
; REFERENCE/DOCKET NUMBER: WU-2855
; TELECOMMUNICATION INFORMATION:
```


TELEPHONE: (314)694-3117
TELEFAX: (314)694-5435
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 173 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-357-820-8

Query Match 75.6%; Score 31; DB 4; Length 173;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
|||||
DB 138 MYPTY 142

RESULT 5
US-09-009-156-5
Sequence 5, Application US/09009156
Patent No. 6046031
GENERAL INFORMATION:
APPLICANT: Ni, Jian
APPLICANT: Ruben, Steven
APPLICANT: Brewer, Laurie
APPLICANT: Gentz, Reiner
TITLE OF INVENTION: No. 6046031el Metalloproteinas
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: US
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/009,156
FILING DATE:

CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/034,205
FILING DATE: 21-JAN-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/049,607
FILING DATE: 13-JUN-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/054,541
FILING DATE: 01-AUG-1997
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PF376
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-309-8504
TELEFAX: 301-309-8439
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 261 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-009-156-5

Query Match 75.6%; Score 31; DB 3; Length 261;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
|||||
DB 226 MYPTY 230

RESULT 6
US-09-009-156-6
Sequence 6, Application US/09009156
Patent No. 6046031
GENERAL INFORMATION:
APPLICANT: Ni, Jian
APPLICANT: Ruben, Steven
APPLICANT: Brewer, Laurie
APPLICANT: Gentz, Reiner
TITLE OF INVENTION: No. 6046031el Metalloproteinas
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: US
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/009,156
FILING DATE:

CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/034,205
FILING DATE: 21-JAN-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/049,607
FILING DATE: 13-JUN-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/054,541
FILING DATE: 01-AUG-1997
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PF376
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-309-8504
TELEFAX: 301-309-8439
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 264 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-009-156-6

Query Match 75.6%; Score 31; DB 3; Length 264;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
|||||
DB 232 MYPTY 236

RESULT 7
US-09-009-156-7

US-08-448-489-18
 ; Sequence 18, Application US/08448489
 ; Patent No. 6184022
 ; GENERAL INFORMATION:
 ; APPLICANT: SEIKI, Motoharu
 ; APPLICANT: SATO, Hiroshi
 ; APPLICANT: SHINAGAWA, Akira
 ; FILE OF INVENTION: NOVEL METALLOPROTEINASE AND ENCODING DNA THEREFOR
 ; FILE REFERENCE: 55-290P
 ; CURRENT APPLICATION NUMBER: US/08/448,489
 ; CURRENT FILING DATE: 1995-06-07
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 18
 ; LENGTH: 267
 ; TYPE: PRT
 ; ORGANISM: Unknown
 ; FEATURE:
 ; OTHER INFORMATION: Description of Unknown Organism: Known Member of
 ; OTHER INFORMATION: Matrix Metalloproteinase Family
 ; US-08-448-489-18

Query Match 75.6%; Score 31; DB 4; Length 267;
 Best Local Similarity 100.0%; Pred. No. 69;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 MYPTY 6
 Db 232 MYPTY 236

RESULT 8
 US-08-896-062-2
 ; Sequence 2, Application US/08896062
 ; Patent No. 6010693
 ; GENERAL INFORMATION:
 ; APPLICANT: KHIRA, Yasunori
 ; TITLE OF INVENTION: PROCESS FOR PRODUCING HUMAN MATRILYSIN
 ; FILE OF INVENTION: BY MEANS OF RECOMBINANT DNA
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Browdy and Neimark
 ; STREET: 419 Seventh Street N.W., Ste. 300
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: United States of America
 ; ZIP: 20004

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/896,062
 FILING DATE: 17-JUL-1997
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/530,984
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Browdy, Roger L.
 REGISTRATION NUMBER: 25,618
 REFERENCE/DOCKET NUMBER: KHIRA=2
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 628-5197
 TELEFAX: (202) 737-3528
 TELEX: 248633
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 271 amino acids
 TYPE: amino acid
 TOPOLOGY: linear

US-08-896-062-2
 ; MOLECULE TYPE: protein

Query Match 75.6%; Score 31; DB 3; Length 271;
 Best Local Similarity 100.0%; Pred. No. 70;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 MYPTY 6
 Db 236 MYPTY 240

RESULT 9
 US-08-630-592-2
 ; Sequence 2, Application US/08630592
 ; Patent No. 5770432
 ; GENERAL INFORMATION:
 ; APPLICANT: Nishina, Betsy
 ; APPLICANT: No. 5770432 Trauth, Konrad
 ; APPLICANT: Naggett, J. Michael
 ; TITLE OF INVENTION: Obesity Associated Genes
 ; NUMBER OF SEQUENCES: 25
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: FLEHR, HOERHACH, TEST, ALBRITTON & HERBERT
 ; STREET: 3400 Embarcadero Center, Suite 3400
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 941114187
 ; COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MSDOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/630,592
 FILING DATE:
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Sherwood, Pamela J.
 REGISTRATION NUMBER: 36,677
 REFERENCE/DOCKET NUMBER: A59504/BIR/PJS
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 7811989
 TELEFAX: (415) 3983249
 TELEX: 910 277299
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 459 amino acids
 TYPE: amino acid
 STEREOBONDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-630-592-2

Query Match 75.8%; Score 31; DB 1; Length 459;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 MYPTY 6
 Db 233 MYPTY 237

RESULT 10
 US-08-714-991-2
 ; Sequence 2, Application US/08714991
 ; Patent No. 5776762
 ; GENERAL INFORMATION:
 ; APPLICANT: NORTH, Michael

APPLICANT: NISHINA, Patsy
APPLICANT: No. 5776762en-Trauth, Konrad
APPLICANT: NAGGERT, Juergen
TITLE OF INVENTION: OBESITY ASSOCIATED GENES
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/714,991
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: SHERWOOD, Pamela J.
REGISTRATION NUMBER: 36,677
REFERENCE/DOCKET NUMBER: A-59504-1/PJS
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-494-8700
TELEFAX: 415-494-8771
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 459 amino acids
TYPE: amino acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-714-991-2

Query Match 75.6%; Score 31; DB 1; Length 459;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTI 6
Db 233 MYPTI 237

SULT 11
US-09-032-365A-2
Sequence 2, Application US/09032365A
Patent No. 6114502
GENERAL INFORMATION:
APPLICANT: Nishina, Patsy
APPLICANT: No. 6114502th, Michael
APPLICANT: NAGGERT, Juergen
APPLICANT: No. 6114502en-Trauth, Konrad
TITLE OF INVENTION: GENE FAMILY ASSOCIATED WITH
TITLE OF INVENTION: NEUROSENSORY DEFECTS
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bozicevic & Reed, LLP
STREET: 285 Hamilton Avenue, Suite 200
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/032,365A
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Sherwood, Pamela J.
REGISTRATION NUMBER: 36,677
REFERENCE/DOCKET NUMBER: SEQ-2CIP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-327-3400
TELEFAX: 650 327-3231
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 459 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-032-365A-2

Query Match 75.6%; Score 31; DB 3; Length 459;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTI 6
Db 233 MYPTI 237

RESULT 12
US-08-630-592-7
Sequence 7, Application US/08630592
Patent No. 5770432
GENERAL INFORMATION:
APPLICANT: Nishina, Patsy
APPLICANT: No. 5770432enTrauth, Konrad
APPLICANT: NAGGERT, Juergen
APPLICANT: No. 5770432th, Michael
TITLE OF INVENTION: Obesity Associated Genes
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
STREET: 3400 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 941114187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/630,592
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sherwood, Pamela J.
REGISTRATION NUMBER: 36,677
REFERENCE/DOCKET NUMBER: A59504/BIR/PJS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 7811989
TELEFAX: (415) 3983249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 460 amino acids
TYPE: amino acid
STRANDEDNESS: double

TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-630-592-7

Query Match 75.6%; Score 31; DB 1; Length 460;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 MYPTY 6
Db 234 MYPTY 238

RESULT 13
US-08-714-991-7
; Sequence 7, Application US/08714991
; Patent No. 5776762
; GENERAL INFORMATION:
; APPLICANT: NORTH Michael
; APPLICANT: NISHINA, Patsy
; APPLICANT: No. 6114502en-Trauth, Konrad
; TITLE OF INVENTION: GENE FAMILY ASSOCIATED WITH
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLHR, HOBBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embroidery Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; APPLICATION DATA: US/08/714,991
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: SHERWOOD, Pamela J.
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: A-59504-1/PJS
; TELEPHONE: 415-494-8700
; TELEFAX: 415-494-8771
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 460 amino acids
; TYPE: amino acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-714-991-7

Query Match 75.6%; Score 31; DB 1; Length 460;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 MYPTY 6
Db 234 MYPTY 238

RESULT 14
US-09-032-365A-8
; Sequence 6, Application US/09032365A
; Patent No. 6114502
; GENERAL INFORMATION:

APPLICANT: No. 6114502th, Michael
APPLICANT: Nishina, Patsy
APPLICANT: Naggart, Juergen
APPLICANT: No. 6114502en-Trauth, Konrad
TITLE OF INVENTION: GENE FAMILY ASSOCIATED WITH
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rozicevic & Reed, LLP
STREET: 385 Hamilton Avenue, Suite 200
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94301

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA: US/09/032,365A
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Sherwood, Pamela J.
REGISTRATION NUMBER: 36,677
REFERENCE/DOCKET NUMBER: SEQ-2CIP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-327-3400
TELEFAX: 650-327-3231
TELEX:

INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 460 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-032-365A-8

Query Match 75.6%; Score 31; DB 3; Length 460;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 MYPTY 6
Db 234 MYPTY 238

RESULT 15
US-08-068-392-3
; Sequence 3, Application US/08068392
; Patent No. 6150152
; GENERAL INFORMATION:
; APPLICANT: Shapiro, Steven M.
; TITLE OF INVENTION: Human Macrophage Metalloproteinase
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scott J. Meyer, Monsanto Co., A3SM
; STREET: 800 N. Lindbergh Blvd.
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63167

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/068,392
FILING DATE: 19930528
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Meyer, Scott J.
REGISTRATION NUMBER: 25375
REFERENCE/DOCKET NUMBER: 07-24(12406)A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314)694-3117
TELEFAX: (314)694-5435
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 462 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-068-392-3

```

```
Query Match 75.6%; Score 31; DB 4; Length 462;
Best Local Similarity 100.0%; Pred. NO. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Caps 0;
```

Search completed: September 4, 2001, 15:51:12
Job time: 195 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2001, 15:52:08 ; Search time 49.08 Seconds
(without alignments)
12.416 Million cell updates/sec

Title: US-09-630-345-3

Perfect score: 41

Sequence: 1 MYPTYLK 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

al number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 68: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

pred, No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	90.2	384	2 I51267	collagenase (EC 3.4.24.-) - bullfrog
2	35	85.4	393	2 B86644	transporter ybfD [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
3	34	82.9	436	2 C69764	4-aminobutyrate am
4	33	80.5	218	2 E86750	hypothetical prote
5	33	80.5	269	2 D69011	hypothetical prote
6	32	78.0	340	1 MMBE5	cell fusion protei
7	32	78.0	349	2 JU0458	peroxidase (EC 1.1
8	32	78.0	593	2 JC4884	organic cation tra
9	32	78.0	1263	2 F44466	DNA-directed RNA p
10	31	75.6	247	2 E85515	hypothetical prote
11	31	75.6	262	2 T29698	hypothetical prote
12	31	75.6	267	1 KCHOM	matrilysin (EC 3.4
13	31	75.6	267	2 A57490	matrilysin (EC 3.4
14	31	75.6	301	2 S57923	SEC14 protein - ye
15	31	75.6	378	2 F86307	hypothetical prote
16	31	75.6	385	2 E86359	Similar to seed ma
17	31	75.6	462	2 A42401	macrophage elastas
18	31	75.6	478	2 A86677	phospho-beta-gluc
19	31	75.6	483	2 JC5743	matrix metalloprot
20	31	75.6	495	1 PLWLB	L1 protein - bovin
21	31	75.6	501	1 PLWLB2	L1 protein - bovin
22	31	75.6	501	1 PLWLEP	L1 protein - Europ
23	31	75.6	502	2 T19708	hypothetical prote
24	31	75.6	503	2 T29175	hypothetical prote
25	31	75.6	505	2 S68518	tub protein, brain
26	31	75.6	511	2 T19496	hypothetical prote
27	31	75.6	513	1 PLWLDP	L1 protein - deer
28	31	75.6	516	2 T15633	hypothetical prote
29	31	75.6	565	2 H69363	hypothetical prote

30	31	75.6	662	2 S42826	probable ATPase -
31	31	75.6	2185	1 GNNYSV	genome polyprotein
32	31	75.6	2185	1 GNNYSV	genome polyprotein
33	30	73.2	141	2 T29507	hypothetical prote
34	30	73.2	152	2 G83476	hypothetical prote
35	30	73.2	155	2 T17986	hypothetical prote
36	30	73.2	192	2 T11930	NADH dehydrogenase
37	30	73.2	239	2 B64757	glycolate oxidase
38	30	73.2	239	2 D85522	probable dehydroge
39	30	73.2	248	2 S57910	probable response
40	30	73.2	262	2 C81384	shikimate 5-dehydr
41	30	73.2	287	2 D83414	probable transcrip
42	30	73.2	343	1 MMBEA5	cell fusion protei
43	30	73.2	368	1 HLBECH	membrane glycoprot
44	30	73.2	410	2 T20397	hypothetical prote
45	30	73.2	491	2 JE0396	phospho-beta-galac

ALIGNMENTS

RESULT 1

I51267
collagenase (EC 3.4.24.-) - bullfrog
C:Species: Rana catesbeiana (bullfrog)
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 22-Jun-1999
C:Accession: I51267
R:Oofusa, K.; Yomori, S.; Yoshizato, K.
Int. J. Dev. Biol. 38, 345-350, 1994
A:Title: Regionally and hormonally regulated expression of genes of collagen and coll
A:Reference number: I51267; MUID:95071832
A:Accession: I51267
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-384 <OOF>
A:Cross-references: GB:S75623; NID:g913070; PIDN:AAB32661.1; PID:g913071
C:Superfamily: interstitial collagenase; hemopexin repeat homology; matrix metallopro
C:Keywords: hydrolase; metalloproteinase; zinc; zymogen
F:53-231/Domain: matrix metalloproteinase homology <MMP>
F:236-381/Domain: hemopexin repeat homology <PNX>
F:81,189,193,199/Binding site: zinc, catalytic (Cys, His, His, His) (inhibited) #stat
F:189,193,199/Binding site: zinc, catalytic (His) (active) #status predicted
F:190/Active site: Glu #status predicted

Query Match 90.2%; Score 37; DB 2; Length 384;
Best Local Similarity 85.7%; Pred. No. 5;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
|||||
DB 207 MYPTYLK 213

RESULT 2

B86644
transporter ybfD [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 23-Mar-2001
C:Accession: B86644
R:Bohlin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarre, K.; Weissenbach, J.; Eh
Genome Res. in press, 2001
A:Title: The complete genome sequence of the lactic acid bacterium.
A:Reference number: A86625
A:Accession: B86644
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-393 <STO>
A:Cross-references: GB:AE005176; NID:g12723004; PIDN:AAK04252.1; GSPDB:GN00146
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: ybfD

Query Match 85.4%; Score 35; DB 2; Length 393;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYL 7
 |||||
 Db 278 MYPTVL 283

RESULT 3
 C:Species: aminobutyrate aminotransferase homolog yncG - Bacillus subtilis
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
 C:Accession: G69764
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertea, C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chodura, A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.; Fure 390, 249-256, 1997
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallier, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Huilo, M.F.; Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon, A.; Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron, A.; Winters, P.; Wipat, A.; Yamamoto, T.; Terpstra, P.; Tojnoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, T.; Terpstra, P.; Tojnoni, A.; Tosato, V.; Uchiyama, A.; Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
 A:Reference number: A6580; MUID:98044033
 A:Accession: C69764
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-436 <XUN>
 A:Cross-references: GB:Z99106; GB:AL009126; NID:G2632653; PIDN:CAB12198.1; PID:G2632691
 A:Experimental source: strain 168
 C:Genetics:
 A:Gene: yncG
 C:Superfamily: ornithine--oxo-acid aminotransferase

Query Match 82.9%; Score 34; DB 2; Length 436;
 Best Local Similarity 71.4%; Pred. No. 23;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYL 8
 |||||
 88 MYPTYIE 94

RESULT 4
 C:Species: hypothetical protein ykdb [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
 C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 23-Mar-2001
 C:Accession: E86750
 R:Boletín, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malmgren, K.; Weissensbach, J.; Ehrlich, A.; Authors: Wincker, P.; Mauger, S.; Jaillon, O.; Malmgren, K.; Weissensbach, J.; Ehrlich, A.; Title: The complete genome sequence of the lactic acid bacterium.
 A:Reference number: A86625
 A:Accession: E86750
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-218 <STO>
 A:Cross-references: GB:AE005176; NID:g12723952; PIDN:AAK05103.1; GSPDB:GN00146
 A:Experimental source: strain IL1403
 C:Genetics:
 A:Gene: ykdb

Query Match 80.5%; Score 33; DB 2; Length 218;

Best Local Similarity 83.3%; Pred. No. 19;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 YPTYLK 8
 |||||
 Db 204 YPTYIK 209

RESULT 5
 C:Species: hypothetical protein MTH1087 - Methanobacterium thermoautotrophicum (strain Delta H)
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
 C:Accession: D69011
 R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, K.; S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
 J. Bacteriol. 179, 7135-7155, 1997
 A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: fu
 A:Reference number: A69000; MUID:98037514
 A:Accession: D69011
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-269 <MTH>
 A:Cross-references: GB:AF000879; GB:AE000666; NID:G2622175; PIDN:AA85576.1; PID:G262
 A:Experimental source: strain Delta H
 C:Genetics:
 A:Gene: MTH1087

Query Match 80.5%; Score 33; DB 2; Length 269;
 Best Local Similarity 83.3%; Pred. No. 23;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 YPTYLK 8
 |||||
 Db 175 YPTYIK 180

RESULT 6
 C:Species: cell fusion protein precursor - human herpesvirus 3
 C:Date: 30-Sep-1988 #sequence_revision 30-Sep-1988 #text_change 16-Jul-1999
 C:Accession: E27212
 R:Davidson, A.J.; Scott, J.E.
 J. Gen. Virol. 67, 1759-1816, 1986
 A:Title: The complete DNA sequence of varicella-zoster virus.
 A:Reference number: A27345; MUID:86306657
 A:Accession: E27212
 A:Molecule type: DNA
 A:Residues: 1-340 <DAV>
 A:Cross-references: EMBL:X04370; NID:G59989; PIDN:CAA27888.1; PID:G59994
 C:Genetics:
 A:Gene: 5
 C:Superfamily: herpesvirus cell fusion protein
 F:Keywords: membrane fusion; transmembrane protein
 F:323/Domain: signal sequence #status predicted <SIG>
 F:32340/Product: cell fusion protein #status predicted <CFP>
 F:115-137/Domain: transmembrane #status predicted <TM1>
 F:220-238/Domain: transmembrane #status predicted <TM3>
 F:251-269/Domain: transmembrane #status predicted <TM4>
 F:307-322/Domain: transmembrane #status predicted <TM5>

Query Match 79.0%; Score 32; DB 1; Length 340;
 Best Local Similarity 57.1%; Pred. No. 47;
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
 |||||
 Db 244 LYPTVIR 250

A:Reference number: S41462; MUID:94252810
A:Accession: S41466
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1263 <PAL>
A:Cross-references: EMBL:X72695; NID:g425255; PIDN:CAA51246.1; PID:g425258
A:Experimental source: DSM 3109
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1993
A:Accession: S41462
A:Molecule type: protein
A:Residues: 1-5,'X',7-23 <PAW>
A:Experimental source: DSM 3109
R:Liaw, D.; Dennis, P.P.
J. Biol. Chem. 267, 22787-22797, 1992
A:Title: The organization and expression of essential transcription translation compo
A:Reference number: A44466; MUID:93054590
A:Accession: F44466
A:Molecule type: DNA
A:Residues: 1-404 <LIA>
A:Cross-references: EMBL:Z11839; NID:g407020; PIDN:CAA77863.1; PID:g408188
A:Note: sequence extracted from NCBI backbone (NCBIP:118059)
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hic
Garrett, M.E.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson,
C.M.
Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome
A:Reference number: A72200; MUID:99287316
A:Accession: B72373
A:Molecule type: DNA
A:Residues: 1-1263 <ARN>
A:Cross-references: GB:AE001724; GB:AE000512; NID:g4980966; PIDN:AAD35543.1; PID:g4980
A:Experimental source: strain MS88
C:Genetics:

```

Query Match          78.0%; Score 32; DB 2; Length 1263;
Best Local Similarity 83.3%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      2  MYPTYL 7
          :|||||
Db      214  LYPTYL 219

RESULT 10
B85515
hypothetical protein Z0321 [imported] - Escherichia coli (strain O157:H7)
C:Species: Escherichia coli
C:date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 31-Mar-2001

```

Query Match 75.6%; Score 31; DB 2; Length 247;
 Best Local Similarity 83.3%; Pred. No. 54;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 A:Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.
 A:Reference number: A85480; MUID:21074935; PMID:11206551
 A:Accession: B85515
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-247 <STO>
 A:Cross-references: GB:AE005174; NID:q12513034; PIDN:AAG54582.1; GSPDB:GNO0145; UWGP:
 A:Experimental source: strain O157:H7, substrain EDL933
 C:Genetics:
 C:Superfamily: hypothetical protein f242a

QY 3 YPTYLK 8
 Db 76 YPSYLK 81

RESULT 11

T29698
 hypothetical protein F31A3.2 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
 C:Accession: T29698
 R:Murray, J.; Le, T.T.
 submitted to the EMBL Data Library, May 1996
 A:Description: The sequence of C. elegans cosmid F31A3.
 A:Reference number: Z20667
 A:Accession: T29698
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-262 <MUR>
 A:Cross-references: EMBL:U58742; PIDN: AAB36855.1; GSPDB: GN00028; CESP: F31A3.2
 A:Experimental source: strain Bristol N2; clone F31A3
 C:Genetics:
 A:Gene: CESP:F31A3.2
 A:Map position: X
 A:Introns: 23/3; 80/2; 155/2

Query Match

Best Local Similarity 75.6%; Score 31; DB 2; Length 262;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
 Db 188 MYPVILE 194

RESULT 12

matrilysin (EC 3.4.24.23) precursor - human
 N:Altrinate names: matrin; matrix metalloproteinase 7 (MMP7); probable metalloproteinase
 N:Constitutes: Promatrilysin
 C:Species: Homo sapiens (man)
 C>Date: 30-Sep-1997 #sequence_revision 30-Sep-1992 #text_change 18-Jun-1999
 C:Accession: P28816; A60539; S24324
 R:Muller, D.; Quasthoff, S.; Giesel, M.C.; Millon-Collard, R.; Abecassis, J.; Breathnach, Biochem. J. 263: 187-192, 1989
 A:Title: The collagenase gene family in humans consists of at least four members.
 A:Reference number: A90339; MUID: 88339885
 A:Accession: P28816
 A:Molecule type: mRNA
 A:Residues: 1-267 <MUL>
 A:Cross-references: EMBL:X07819; NID: g35798; PIDN: CAA30678.1; PID: g35799
 R: Miyazaki, K.; Hattori, Y.; Umenishi, F.; Yasumitsu, H.; Umeda, M.
 Cancer Res. 50: 7758-7764, 1990
 A:Title: Purification and characterization of extracellular matrix-degrading metallopro
 A:Reference number: A60539; MUID: 91070531
 A:Accession: A60539
 A:Molecule type: protein
 A:Residues: 18-35, X', 37-42 <MIY>
 R: Marti, H.P.; McNeil, L.; Thomas, G.; Davies, M.; Lovett, D.H.
 Biochem. J. 285: 899-905, 1992
 A:Title: Molecular characterization of a low-molecular-mass matrix metalloproteinase sec
 A:Reference number: S24324; MUID: 92359961
 A:Accession: S24324
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-267 <MAR>
 A:Cross-references: EMBL: Z11887; NID: g35802; PIDN: CAA77942.1; PID: g35803
 C:Comment: This enzyme is similar in its activity to stromelysin and degrades various ex
 s of types II, IV, IX, X, and XI.
 C:Comment: Matrilysin hydrolyzes peptide bonds in plasminogen to yield a fragment with a
 C:Genetics:

A:Gene: GDB:MMP7; MPSL1
 A:Cross-references: GDB:125751; OMIM:178990
 A:Map position: 11q21-11q22
 C:Superfamily: matrilysin; matrix metalloproteinase homology
 C:Keywords: calcium; extracellular matrix; fibroblast; hydrolase; metalloproteinase;
 F:1-17/Domain: signal sequence #status predicted <SIG>
 F:18-267/Product: promatrilysin #status predicted <PRO>
 F:18-94/Domain: activation peptide #status predicted <ACT>
 F:55-259/Domain: activation peptide #status predicted <ACT>
 F:85-92/Region: autolysis site
 F:95-267/Product: matrilysin #status predicted <MAT>
 F:87-214, 218, 224/Binding site: zinc, catalytic (Cys, His, His) (inhibited) #stat
 F:214, 218, 224/Binding site: zinc, catalytic (His) (active) #status predicted
 F:215/Active site: Glu #status predicted

Query Match 75.6%; Score 31; DB 1; Length 267;
 Best Local Similarity 100.0%; Pred. No. 59;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
 Db 232 MYPTY 236

RESULT 13

A57490
 matrilysin (EC 3.4.24.23) precursor - rat
 N:Alternate names: matrix metalloproteinase 7 (MMP7)
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 08-Dec-1995 #sequence_revision 08-Dec-1995 #text_change 22-Jun-1999
 C:Accession: A57490
 R: Abramson, S.R.; Conner, G.E.; Nagase, H.; Neuhaus, I.; Woessner Jr., J.F.
 J. Biol. Chem. 270: 16016-16022, 1995
 A:Title: Characterization of rat uterine matrilysin and its cDNA. Relationship to hum
 A:Reference number: A57490; MUID: 95332299
 A:Accession: A57490
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-267 <ABR>
 A:Cross-references: GB: I24374; NID: g402492; PIDN: AAA99432.1; PID: g402493
 C:Superfamily: matrilysin; matrix metalloproteinase homology
 C:Keywords: hydrolase; metalloproteinase; zinc; zymogen
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-267/Product: matrilysin #status predicted <MAT>
 F:58-262/Domain: matrix metalloproteinase homology <MMP>
 F:90, 217, 221, 227/Binding site: zinc, catalytic (Cys, His, His) (inhibited) #stat
 F:217, 221, 227/Binding site: zinc, catalytic (His) (active) #status predicted
 F:218/Active site: Glu #status predicted

Query Match 75.6%; Score 31; DB 2; Length 267;
 Best Local Similarity 100.0%; Pred. No. 59;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
 Db 235 MYPTY 239

RESULT 14

S57923
 SEC14 protein - yeast (Candida albicans)
 C:Species: Candida albicans
 C>Date: 13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 26-Feb-1998
 C:Accession: S57923; S72193
 R: Montelliva, L.; Sanchez, M.; Pla, J.; Gil, C.; Nombela, C.
 Submitted to the EMBL Data Library, September 1994
 A:Description: Characterisation of the Candida albicans SEC14 homolog gene.
 A:Reference number: S57923
 A:Accession: S57923
 A:Status: preliminary
 A:Molecule type: DNA

A;Residues: 1-301 <MON>
 A;Cross-references: EMBL:X81937
 R;Monteoliva, L.; Sanchez, M.; Pla, J.; Gil, C.; Nombela, C.
 Yeast 12, 1097-1105, 1996
 A;Title: Cloning of Candida albicans SEC14 gene homologue coding for a putative essential protein
 A;Reference number: S72193; MUID:97051600
 A;Accession: S72193
 A;Molecule type: DNA
 A;Residues: 1-301 <NOW>
 A;Cross-references: EMBL:X81937
 A;Note: the authors translated the codon CTG for residue 180 as Ser
 C;Genetics:
 A;Gene: SEC14
 C;Superfamily: cellular retinaldehyde-binding protein; cellular retinaldehyde-binding protein
 F;59-266/Domain: cellular retinaldehyde-binding protein homology <CRB>

Query Match 75.6%; Score 31; DB 2; Length 301;
 Best Local Similarity 100.0%; Pred. No. 66;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2 MYPTY 6
 |||||
 Db 106 MYPTY 110

RESULT 15
 F86307
 hypothetical protein AAD50017.1 [imported] - Arabidopsis thaliana
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C;Accession: F86307
 R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 Jensen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A;Reference number: A86141; MUID:21016719
 A;Accession: F86307
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-378 <STO>
 A;Cross-references: GB:AE005172; NID:g5734752; PIDN:AAD50017.1; GSPDB:GN00141
 C;Genetics:
 Map position: 1

Query Match 75.6%; Score 31; DB 2; Length 378;
 Best Local Similarity 100.0%; Pred. No. 84;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
 |||||
 Db 118 MYPTY 122

Search completed: September 4, 2001, 15:52:08
 Job time: 231 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2001, 15:54:09 ; Search time 27.19 Seconds
(without alignments)
10.079 Million cell updates/sec

Title: US-09-630-345-3
Perfect score: 41
Sequence: 1 XMYPTYLK 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues
Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37	90.2	384	MM01_RANCA	Q11133 rana catesb
2	34	82.9	436	GABT_BAGSU	P94427 bacillus su
3	32	78.0	340	CELF_VZVD	P09261 varicella-z
4	32	78.0	349	PERE_ARATH	P24102 arabidopsis
5	32	78.0	594	NU5M_HIPAM	Q92zy1 hippopotamu
6	32	78.0	1263	RPOB_THEMA	P29398 thermotoga
7	31	75.6	262	MM07_FELCA	P55032 felis silve
8	31	75.6	264	MM07_MOUSE	Q10738 mus musculu
9	31	75.6	267	MM07_HUMAN	P09237 homo sapien
10	31	75.6	267	MM07_RAT	P50280 rattus norv
11	31	75.6	301	SC14_CANAL	P46250 candida alb
12	31	75.6	462	MM12_MOUSE	P34960 mus musculu
13	31	75.6	465	MM12_RAT	O63341 rattus norv
14	31	75.6	467	MM18_XENLA	O13065 xenopus lae
15	31	75.6	482	MM20_MOUSE	P57748 mus musculu
16	31	75.6	483	MM20_HUMAN	O60882 homo sapien
17	31	75.6	483	MM20_PIG	P79287 sus scrofa
18	31	75.6	495	VLL1_BPVI	P03103 bovine papi
19	31	75.6	497	VLL1_BPVI	P06458 bovine papi
20	31	75.6	501	VLL1_PAPVE	P11326 european el
21	31	75.6	503	YPTJ_CAEEL	Q23469 caenorhabd
22	31	75.6	505	TUB_MOUSE	P50586 mus musculu
23	31	75.6	505	TUB_RAT	O88808 rattus norv
24	31	75.6	506	TUB_HUMAN	P50607 homo sapien
25	31	75.6	511	VELJ_CAEEL	P90756 caenorhabd
26	31	75.6	513	VLL1_PAPVD	P03104 deer papill
27	31	75.6	662	YWEJ_SCHMA	P46508 schistosoma
28	31	75.6	2185	POLG_SVDVH	P16604 s genome po
29	31	75.6	2185	POLG_SVDVU	P13900 s genome po
30	30	73.2	91	Y13F_BPVI	P39499 bacterioph
31	30	73.2	192	NUGM_PROVI	P37622 prototheca
32	30	73.2	239	YKGE_ECOLI	P77252 escherichia
33	30	73.2	308	Y04O_BPVI	P39254 bacterioph

RESULT 1	MM01_RANCA	STANDARD;	PRT;	384 AA.
ID	MM01_RANCA			
AC	Q11133;			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	01-OCT-2000 (Rel. 40, Last annotation update)			
DE	INTERSTITIAL COLLAGENASE PRECURSOR (EC 3.4.24.7) (MATRIX METALLOPROTEINASE-1) (MMP-1) (TC1).			
OS	Rana catesbeiana (Bull frog).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Neobatrachia; Ranidae; Rana.			
OX	NCBI_TaxID=8400;			
NP	[1]			
RN	SEQUENCE FROM N. A.			
RC	TISSUE=Skin;			
RX	MEDLINE=95071832; PubMed=7981043;			
RA	Oofusa K., Yomori S., Yoshizato K.;			
RT	"Regionally and hormonally regulated expression of genes of collagen and collagenase in the anuran larval skin.";			
RL	Int. J. Dev. Biol. 38:345-350(1994).			
CC	-/- FUNCTION: CLEAVES COLLAGENS OF TYPES I, II, AND III AT ONE SITE IN THE HELICAL DOMAIN. ALSO CLEAVES COLLAGENS OF TYPES VII AND X.			
CC	-/- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY.			
CC	-/- ENZYME REGULATION: CAN BE ACTIVATED WITHOUT REMOVAL OF THE ACTIVATION PEPTIDE (BY SIMILARITY).			
CC	-/- SIMILARITY: CONTAINS 1 HEMOPEXIN-LIKE DOMAIN.			
CC	-/- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC METALLOPROTEASE) ALSO KNOWN AS MATRININ SUBFAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; S75623; AAB32661.1;			
DR	HSP; P21692; IFBL.			
DR	MEROPS; M10.001;			
DR	InterPro; IPR000130;			
DR	InterPro; IPR000585;			
DR	InterPro; IPR001816;			
DR	Pfam; PF00413; Peptidase_M10; 1.			
DR	Pfam; PF00045; hemopexin; 2.			
DR	PROSITE; PS00024; HEMOPEXIN; FALSE_NEG.			
DR	PROSITE; PS00142; ZINC_PROTEASE; 1.			
DR	PROSITE; PS00546; CYSTEINE_SWITCH; FALSE_NEG.			
KW	Hydrolase; Metalloprotease; zinc; Zymogen; Calcium; Collagen degradation; Extracellular matrix; Signal.			
FT	SIGNAL 1 25 POTENTIAL.			
FT	PROPEP 26 88 ACTIVATION PEPTIDE (POTENTIAL).			
FT	CHAIN 89 384 INTERSTITIAL COLLAGENASE.			
FT	DOMAIN 239 384 HEMOPEXIN-LIKE.			
FT	SITE 81 81 CYSTEINE SWITCH (POTENTIAL).			

P28933 equine herp
P08560 human cytom
P36568 serratia ma
P15362 mycoplasma
P21042 vaccinia vi
P17372 vaccinia vi
P03261 human adeno
P04495 human adeno
Q03749 bacillus th
Q45707 bacillus th
Q45708 bacillus th
P06578 bacillus th

```

FT METAL 189 189 ZINC (CATALYTIC) (BY SIMILARITY).
FT AC1_SITE 190 190 BY SIMILARITY.
FT METAL 193 193 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 199 199 ZINC (CATALYTIC) (BY SIMILARITY).
FT DISULFID 242 381 PROBABLE.
SQ SEQUENCE 384 AA; 43582 MW; A5B5E2FB332239DF CRC64;

Query Match 90.2%; Score 37; DB 1; Length 384;
Best Local Similarity 85.7%; Pred. No. 2.9;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
DB 207 MYPTYLK 213
|||||:

RESULT 2
GABT_BACSU STANDARD; PRT; 436 AA.
P94427;
15-DEC-1998 (Rel. 37, Created)
15-DEC-1998 (Rel. 37, Last sequence update)
30-MAY-2000 (Rel. 39, Last annotation update)
DE PROBABLE 4-AMINOBTYRATE AMINOTRANSFERASE (EC 2.6.1.19) (GAMMA-AMINO-
DE N-BUTYRATE TRANSAMINASE) (GABA TRANSAMINASE) (GLUTAMATE-SUCCINIC
DE SEMIALDEHYDE TRANSAMINASE) (GABA AMINOTRANSFERASE) (GABA-AT).
GN GABT.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
EX MEDLINE=97124189; PubMed=8969502;
RA Yamane K., Kumano M., Kurita K.;
RT "The 25 degrees-36 degrees region of the Bacillus subtilis
RT chromosome: determination of the sequence of a 146 kb segment and
RT identification of 113 genes.";
RL Microbiology 142:3047-3056(1996).
CC -1- CATALYTIC ACTIVITY: 4-AMINOBTYRANATE + 2-OXOGLUTARATE = SUCCINATE
CC -1- SEMIALDEHYDE + L-GLUTAMATE.
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -1- PATHWAY: 4-AMINOBTYRATE (GABA) DEGRADATION PATHWAY.
CC -1- SIMILARITY: BELONGS TO CLASS-III OF PYRIDOXAL-PHOSPHATE-DEPENDENT
CC AMINOTRANSFERASES.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; D50453; BAA09021.1; -
DR EMBL; Z99105; CAB12198.1; -
DR SUBTILIST; BGL1043; gabt.
DR InterPro; IPR000994; -
DR Pfam; PF00202; aminotran.3; 1.
DR PROSITE; PS00600; AA_TRANSFER_CLASS.3; 1.
KW Transferase; Amino transferase; Pyridoxal phosphate.
FT BINDING 281 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ SEQUENCE 436 AA; 47249 MW; D0961F6D4189A8F3 CRC64;

Query Match 82.9%; Score 34; DB 1; Length 436;
Best Local Similarity 71.4%; Pred. No. 13;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
|||||:

```

```

Db 88 MYPTYIE 94

RESULT 3
ID CELF_VZVD STANDARD; PRT; 340 AA.
AC P09261;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE CELL FUSION PROTEIN PRECURSOR.
GN 5.
OS Varicella-zoster virus (strain Dumas) (VZV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10338;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=86306657; PubMed=3018124;
RA Davison A.J., Scott J.E.;
RT "The complete DNA sequence of varicella-zoster virus.";
RL J. Gen. Virol. 67:1759-1816(1986).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X04370; CAA27888.1; -
DR PIR; E27212; MMBE5
DR InterPro; IPR002567; -
DR Pfam; PF01621; Fusion_gly_K; 1.
KW Fusion protein; Transmembrane; Signal.
FT SIGNAL 1 ?
FT CHAIN 1 ? 340 CELL FUSION PROTEIN.
FT CARBOHYD 55 55 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 65 65 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 340 AA; 38576 MW; 0387FE00EC39C946 CRC64;

Query Match 78.0%; Score 32; DB 1; Length 340;
Best Local Similarity 57.1%; Pred. No. 25;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
DB 244 LYPTVIR 250
|||||:

RESULT 4
ID PERE_ARATH STANDARD; PRT; 349 AA.
AC P24102;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE BASIC PEROXIDASE E PRECURSOR (EC 1.11.1.7).
GN PRXEA.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=91200671; PubMed=2016063;
RA Intapurk C., Higashimura N., Yamamoto K., Okada N., Shimmyo A.,
RA Takano M.;
RT "Nucleotide sequences of two genomic DNAs encoding peroxidase of
RT Arabidopsis thaliana.";

```

RL Gene 98:237-241(1991).

CC -1- FUNCTION: REMOVAL OF H(2)O(2), OXIDATION OF TOXIC REDUCTANTS,

CC BIOSYNTHESIS AND DEGRADATION OF LIGNIN, DEFENSE RESPONSE TOWARD

CC WOUNDING ON EACH ISOZYME/ISOFORM. THESE FUNCTIONS MIGHT BE

CC DEPENDENT ON METABOLISM OF AUXIN. IN EACH PLANT TISSUE.

CC -1- CATALYTIC ACTIVITY: DONOR + H(2)O(2) = OXIDIZED DONOR + 2 H(2)O.

CC -1- COFACTOR: HEME.

CC -1- TISSUE SPECIFICITY: ROOTS.

CC -1- FAMILY: BELONGS TO THE PEROXIDASE FAMILY. PLANT PEROXIDASE

CC SUBFAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; M58381; AAA32842.1; -.

CC PIR; J00458; J00458.

CC HSSP; P00433; LATJ.

CC InterPro; IPR000823; -.

CC DR InterPro; IPR002016; -.

CC DR Pfam; PF00141; peroxidase.1.

CC DR PRINTS; PR00458; PEROXIDASE.

CC DR PRINTS; PR00461; PLPEROXIDASE.

CC DR PROSITE; PS00435; PEROXIDASE.1; 1.

CC DR PROSITE; PS00436; PEROXIDASE.2; 1.

CC KW Oxidoreductase; Glycoprotein; Peroxidase; Heme; Multigene family;

CC KW Signal.

FT SIGNAL 1 19

FT CHAIN 20 349 BASIC PEROXIDASE E.

FT ACT_SITE 67 67 BY SIMILARITY.

FT ACT_SITE 71 71 DISTAL HISTIDINE (BY SIMILARITY).

FT ACT_SITE 199 199 PROXIMAL HISTIDINE (HEME AXIAL LIGAND).

FT DISULFID 40 120 BY SIMILARITY.

FT DISULFID 73 78 BY SIMILARITY.

FT DISULFID 126 329 BY SIMILARITY.

FT DISULFID 206 238 BY SIMILARITY.

FT CARBOHYD 86 86 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 173 173 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 187 187 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 217 217 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 243 243 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC SEQUENCE 349 AA; 6E02C5758C36A8BF CRC64;

Query Match 78.0%; Score 32; DB 1; Length 349;

Best Local Similarity 83.3%; Pred. No. 26;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYL 7

Db 226 LYPTYL 231

RESULT 5

ID NU5M_HIPAM STANDARD; PRT; 594 AA.

AC Q92X1;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 5 (EC 1.6.5.3).

GN MN5D5 OR ND5 OR NADH5.

OS Hippopotamus amphibius (Hippopotamus).

OC Mitochondrion.

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Hippopotamidae; Hippopotamus.

OX NCBI_TaxID=9833;

RN [1]

RP SEQUENCE FROM N.A.

RA Ursing B.M., Arnason U.;

RT "Analyses of mitochondrial genomes strongly support a hippopotamus-

RT whale clade.";

RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.

CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; AJ010957; CA009438.1; -.

CC InterPro; IPR001516; -.

CC DR InterPro; IPR001750; -.

CC DR Pfam; PF00361; oxidored_ql.1.

CC DR Pfam; PF00662; oxidored_ql_N.1.

CC KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.

CC SEQUENCE 594 AA; 66599 MW; 4130499096B5A5CE CRC64;

Query Match 78.0%; Score 32; DB 1; Length 594;

Best Local Similarity 83.3%; Pred. No. 44;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 YPTVLK 8

Db 32 YPTVVK 37

RESULT 6

ID RPOB_THEMA STANDARD; PRT; 1263 AA.

AC P29358;

DT 01-DEC-1992 (Rel. 24, Created)

DT 01-JUN-1994 (Rel. 29, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE DNA-DIRECTED RNA POLYMERASE BETA CHAIN (EC 2.7.7.6) (TRANSCRIPTASE

DE BETA CHAIN) (RNA POLYMERASE BETA SUBUNIT).

GN RPOB OR TM0458

OS Thermotoga maritima.

OC Bacteria; Thermotogales; Thermotoga.

OX NCBI_TaxID=2336;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=MSB8 / DSM 3109;

RX MEDLINE=94232816; PubMed=8177738;

RA Palm P., Schleper C., Arnold-Ammer I., Holz I., Meier T.,

RA Lottspeich F., Zillig W.;

RT "The DNA-dependent RNA-polymerase of Thermotoga maritima;

RT characterisation of the enzyme and the DNA-sequence of the genes for

RT the large subunits.";

RL Nucleic Acids Res. 21:4904-4908(1993).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=MSB8 / DSM 3109;

RX MEDLINE=99287316; PubMed=10360571;

RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,

RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,

RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,

RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,

RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,

RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;

RT "Evidence for lateral gene transfer between Archaea and Bacteria from

RT genome sequence of Thermotoga maritima.";

RL Nature 399:323-329(1999).

RN [3]

RP SEQUENCE OF 1-404 FROM N.A.

RC STRAIN=MSB8 / DSM 3109;

RX MEDLINE=93054590; PubMed=1429627;

RN Liao D., Dennis P.P.;

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

CC EMBL; L36238; AAA99984.1; JOINED.
 CC EMBL; L36243; AAA99984.1; JOINED.
 CC EMBL; L36242; AAA99984.1; JOINED.
 CC EMBL; L36241; AAA99984.1; JOINED.
 CC EMBL; L36240; AAA99984.1; JOINED.
 CC EMBL; L36239; AAA99984.1; JOINED.
 CC EMBL; L36244; AAA99984.1; JOINED.
 CC HSP; P09237; IMMR.
 CC MEROPS; M10.008; --
 CC MGD; MGI:103189; Mmp7.
 CC InterPro; IPR000130; --
 CC InterPro; IPR001818; --
 CC Pfam; PF00413; Peptidase_M10; 1.
 CC PRINTS; PR00138; MATRIXIN.
 CC PROSITE; PS00142; ZINC_PROTEASE; 1.
 CC PROSITE; PS00546; CYSTEINE_SWITCH; 1.
 CC Hydrolyase; Metalloprotease; Zinc; Zymogen; Calcium;
 KW Collagen degradation; Extracellular matrix; Signal.
 FT SIGNAL 1 17 POTENTIAL.
 FT PROPEP 18 94 ACTIVATION PEPTIDE (BY SIMILARITY).
 FT CHAIN 95 264 MATRILYSIN.
 FT SITE 87 87 CYSTEINE SWITCH (POTENTIAL).
 FT METAL 214 214 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACT_SITE 215 215 BY SIMILARITY.
 FT METAL 218 218 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 224 224 ZINC (CATALYTIC) (BY SIMILARITY).
 FT CONFLICT 201 201 G -> D (IN AAA99983).
 SQ SEQUENCE 264 AA; 29755 MW; EDA31A5EBAC63342 CRC64;

Query Match 75.6%; Score 31; DB 1; Length 264;
 Best Local Similarity 100.0%; Pred. No. 31;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTI 6
 Db 232 MYPTI 236
 |||||

RESULT 9
 MM07_HUMAN STANDARD; PRT; 267 AA.
 P09237;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE MATRILYSIN PRECURSOR (EC 3.4.24.23) (PUMP-1 PROTEASE) (UTERINE
 DE METALLOPROTEINASE) (MATRIX METALLOPROTEINASE-7) (MMP-7) (MATRIN).
 GN MMP7 OR MPSL1 OR PUMPL.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88339885; PubMed=2844164;
 RA Muller D., Quantin B., Gesnel M.-C., Millon-Collard R., Abecassis J.,
 RA Breathnach R.;
 RT "The collagenase gene family in humans consists of at least four
 RT members.";
 RL Biochem. J. 253:187-192(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=92359961; PubMed=1497627;
 RA Marti H.P., McNeil L., Thomas G., Davies M., Lovett D.H.;

RT "Molecular characterization of a low-molecular-mass matrix
 RL metalloproteinase secreted by glomerular mesangial cells as PUMP-1.";
 RL Biochem. J. 285:899-905(1992).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=94124555; PubMed=8294454;
 RA Gaire M., Magbanua Z., McDonnell S., McNeil L.B., Lovett D.H.,
 RA Matrisian L.M.;
 RT "Structure and expression of the human gene for the matrix
 RT metalloproteinase matrilysin.";
 RL J. Biol. Chem. 269:2032-2040(1994).
 RN [4]
 RP SEQUENCE OF 18-42.
 RX MEDLINE=91070531; PubMed=2253219;
 RA Miyazaki K., Hattori Y., Umenishi F., Yasumitsu H., Umeda M.;
 RT "Purification and characterization of extracellular matrix-degrading
 RT metalloproteinase, matrin (pump-1), secreted from human rectal
 RT carcinoma cell line.";
 RL Cancer Res. 50:7758-7764(1990).
 RN [5]
 RP FUNCTION.
 RX MEDLINE=89375247; PubMed=2550050;
 RA Quantin B., Murphy G., Breathnach R.;
 RT "Pump-1 cDNA codes for a protein with characteristics similar to
 RT those of classical collagenase family members.";
 RL Biochemistry 28:5327-5334(1989).
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
 RX MEDLINE=95275856; PubMed=7756291;
 RA Browner M.F., Smith W.W., Castelhan A.L.;
 RT "Matrilysin-inhibitor complexes: common themes among
 RT metalloproteases.";
 RL Biochemistry 34:6602-6610(1995).
 CC -1- FUNCTION: DEGRADERS CASEIN, GELATINS OF TYPES I, III, IV, AND V,
 CC AND FIBRONECTIN. ACTIVATES PROCOLLAGENASE.
 CC -1- CATALYTIC ACTIVITY: CLEAVAGE OF 14-ALA-1-LEU-15 AND 16-TYR-1-
 CC LEU-17 IN B CHAIN OF INSULIN. NO ACTION ON COLLAGEN TYPES I, II,
 CC IV, V. CLEAVES GELATIN CHAIN ALPHA-2(I) > ALPHA-1(I).
 CC -1- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC
 CC METALLOPROTEASE) ALSO KNOWN AS MATRININ SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X07819; CAA30678.1; --
 CC EMBL; Z11887; CAA77942.1; --
 CC EMBL; L22524; AAC37543.1; --
 CC EMBL; L22519; AAC37543.1; JOINED.
 CC EMBL; L22520; AAC37543.1; JOINED.
 CC EMBL; L22521; AAC37543.1; JOINED.
 CC EMBL; L22522; AAC37543.1; JOINED.
 CC EMBL; L22523; AAC37543.1; JOINED.
 CC PIR; B28816; KCHUM.
 CC PIR; S24324; S24324.
 CC PDB; IMMP; 03-APR-96.
 CC PDB; IMMQ; 03-APR-96.
 CC PDB; IMMR; 03-APR-96.
 CC MEROPS; M10.008; --
 CC MIM; I78990; --
 CC InterPro; IPR000130; --
 CC InterPro; IPR001818; --
 CC Pfam; PF00413; Peptidase_M10; 1.
 CC PRINTS; PR00138; MATRIXIN.
 CC PROSITE; PS00142; ZINC_PROTEASE; 1.
 CC PROSITE; PS00546; CYSTEINE_SWITCH; 1.
 KW Hydrolyase; Metalloprotease; Zinc; Zymogen; Calcium; 3D-structure;

KW Collagen degradation; Extracellular matrix; Signal; Polymorphism.

FT SIGNAL 1 17 64
FT PROPEP 18 94
FT CHAIN 95 267
FT SITE 87 87
FT METAL 214 214
FT ACT_SITE 215 215
FT METAL 218 218
FT METAL 224 224
FT VARIANT 77 77
R -> H.
/FTID=VAR_006729.

SQ SEQUENCE 267 AA; 29677 MW; F6BDIFC0ADA23603 CRC64;

Query Match 75.6%; Score 31; DB 1; Length 267;

Best Local Similarity 100.0%; Pred. No. 32;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 MYPTY 6

|||||

232 MYPTY 236

RESULT 10

MM07_RAT ID MM07_RAT STANDARD; PRT; 267 AA.

AC P50280;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE MATRILYSIN PRECURSOR (EC 3.4.24.23) (PUMP-1 PROTEASE) (UTERINE

DE METALLOPROTEINASE) (MATRIX METALLOPROTEINASE-7) (MMP-7) (MATRIN).

GN MMP7 OR MMP-7.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=SPRAGUE-DAWLEY; TISSUE=uterus;

EX MEDLINE=95332299; PubMed=7608162.

RA Abramson S.R., Conner G.E., Ragase H., Neuhaus I., Woessner J.F.;

RT "Characterization of rat uterine matrilysin and its cDNA.

RL J. Biol. Chem. 270:16016-16022(1995).

-1- FUNCTION: DEGRADATES CASEIN, GELATINS OF TYPES I, III, IV, AND V,

AND FIBRONECTIN. ACTIVATES PROCOLLAGENASE (BY SIMILARITY).

-1- CATALYTIC ACTIVITY: CLEAVAGE OF 14-ALA-1-LEU-15 AND 16-TYR-1-

LEU-17 IN B CHAIN OF INSULIN. NO ACTION ON COLLAGEN TYPES I, II,

IV, V. CLEAVES GELATIN CHAIN ALPHA-2(I) > ALPHA-1(I).

-1- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY.

-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC

METALLOPROTEASE) ALSO KNOWN AS MATRININ SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; L24374; AAA99432.1; -

CC HSP; P09237; 1MR.

CC MEROPS; M10.008; -

CC InterPro; IPR000130; -

CC InterPro; IPR001818; -

CC Pfam; PF00413; Peptidase_M10; 1.

CC PRINTS; PR00138; MATRININ.

CC PROSITE; PS00142; ZINC_PROTEASE; 1.

CC PROSITE; PS00546; CYSTEINE_SWITCH; 1.

CC Hydrolase; Metalloprotease; zinc; Zymogen; Calcium;

KW Collagen degradation; Extracellular matrix; Signal.

FT SIGNAL 1 20
FT PROPEP 21 97
FT CHAIN 98 267
FT SITE 90 90
FT METAL 217 217
FT ACT_SITE 218 218
FT METAL 221 221
FT METAL 227 227
R -> H.
/FTID=VAR_006729.

SQ SEQUENCE 267 AA; 29885 MW; EBA3C3D9527A4C7B CRC64;

Query Match 75.6%; Score 31; DB 1; Length 267;

Best Local Similarity 100.0%; Pred. No. 32;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 MYPTY 6

|||||

235 MYPTY 239

RESULT 11

SC14_CANAL

ID SC14_CANAL STANDARD; PRT; 301 AA.

AC P46250;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 30-NOV-2000 (Rel. 39, Last annotation update)

DE SEC14 CYTOSOLIC FACTOR (PHOSPHATIDYLINOSITOL/PHOSPHATIDYLCHOLINE

DE TRANSFER PROTEIN) (PI/PC TP).

GN SEC14.

OS Candida albicans (Yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Mitosporic Saccharomycetales; Candida.

OX NCBI_TaxID=5476;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 64385 / 1001;

EX MEDLINE=97051600; PubMed=8896277;

RA Monteleone L., Sanchez M., Pla J., Gil C., Nombela C.;

RT "Cloning of Candida albicans SEC14 gene homologue coding for a

RL putative essential function."

RL Yeast 12:1097-1103(1996).

[2]

RP SEQUENCE FROM N.A.

RC STRAIN=SGY243;

EX MEDLINE=98048477; PubMed=9387231;

RA Riggle P.J., Slobodkin I.V., Brown D.H. Jr., Hanson M.P.,

RT "Two transcripts, differing at their 3' ends, are produced from the

RT Candida albicans SEC14 gene."

RL Microbiology 143:3527-3535(1997).

-1- FUNCTION: REQUIRED FOR TRANSPORT OF SECRETORY PROTEINS FROM THE

CC GOLGI COMPLEX. CATALYZES THE TRANSFER OF PHOSPHATIDYLINOSITOL AND

CC PHOSPHATIDYLCHOLINE BETWEEN MEMBRANES IN VITRO (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: ASSOCIATED WITH THE GOLGI COMPLEX AS A

CC PERIPHERAL MEMBRANE PROTEIN (BY SIMILARITY)

CC -1- SIMILARITY: BELONGS TO THE SEC14 CYTOSOLIC FACTOR FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; X81937; CAA57490.1; -

CC HSP; U61975; AAB41491.1; -

CC HSP; P24280; LAUA.

CC InterPro; IPR001251; -

CC Pfam; PF00650; CRAL_TRIO; 1.

CC Transport; Protein transport; Golgi stack

KW Transport; Protein transport; Golgi stack

SQ SEQUENCE 301 AA; 34709 MW; 1F194A8EA8B525BE CRC64;

Query Match 75.6%; Score 31; DB 1; Length 301;
 Best Local Similarity 100.0%; Pred. No. 36;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPT 6
 DB 106 MYPT 110

RESULT 12
 MM12_MOUSE STANDARD; PRT; 462 AA.
 AC P34960;
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE MACROPHAGE METALLOELASTASE PRECURSOR (EC 3.4.24.65) (MME) (MATRIX
 METALLOPROTEINASE-12) (MMP-12).
 MM12 OR MMEL OR MME.
 Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 99-125.
 RC TISSUE=Macrophage;
 RX MDL=92165826; PubMed=1537850;
 RA Shapiro S.D., Griffin G.L., Gilbert D.J., Jenkins N.A.,
 RA Copeland N.G., Welgus H.G., Senior R.M., Ley T.J.;
 RT "Molecular cloning, chromosomal localization, and bacterial
 expression of a murine macrophage metalloelastase.";
 RL J. Biol. Chem. 267:4664-4671(1992).
 CC -!- FUNCTION: MAY BE INVOLVED IN TISSUE INJURY AND REMODELING. HAS
 CC SIGNIFICANT ELASTOLYTIC ACTIVITY.
 CC -!- CATALYTIC ACTIVITY: HYDROLYSIS OF SOLUBLE AND INSOLUBLE ELASTIN.
 CC SPECIFIC CLEAVAGES ARE ALSO PRODUCED AT 14-ALA-|-LEU-15 AND 16-
 CC TYR-|-LEU-17 IN THE B CHAIN OF INSULIN.
 CC -!- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY.
 CC -!- SIMILARITY: CONTAINS 1 HEMOPEXIN-LIKE DOMAIN.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC
 CC METALLOPROTEASE) ALSO KNOWN AS MATRININ SUBFAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M82831; AAA39526.1; -
 CC PIR; A42401; A42401.
 CC HSP; P03956; ICGL.
 CC MEROPS; M10.009; -
 CC MGD; MGI:97005; Mmp12.
 CC InterPro; IPR000130; -
 CC InterPro; IPR000585; -
 CC InterPro; IPR001818; -
 CC Pfam; PF000413; Peptidase_M10; 1.
 CC Pfam; PF00045; hemopexin; 4.
 CC PRINTS; PR00138; MATRININ.
 CC PROSITE; PS00024; HEMOPEXIN; 1.
 CC PROSITE; PS00142; ZINC_PROTEASE; 1.
 CC PROSITE; PS00546; CYSTEINE_SWITCH; 1.
 CC Hydrolase; Metalloprotease; Glycoprotein; zinc; zymogen; Calcium;
 KW Extracellular matrix; Signal.
 FT SIGNAL 1 17 PROBABLE.
 FT PROPP 18 98 ACTIVATION PEPTIDE.
 FT CHAIN 99 462 MACROPHAGE METALLOELASTASE.
 FT DOMAIN 272 462 HEMOPEXIN-LIKE.
 FT SITE 85 85 CYSTEINE SWITCH (BY SIMILARITY).

FT METAL 211 211 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACT_SITE 212 212 BY SIMILARITY.
 FT METAL 215 215 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 221 221 ZINC (CATALYTIC) (BY SIMILARITY).
 FT CARBOHYD 21 21 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 74 74 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 310 310 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT DISULFID 275 462 BY SIMILARITY.
 SQ SEQUENCE 462 AA; 53841 MW; BB9625906FIDBEDF CRC64;
 Query Match 75.6%; Score 31; DB 1; Length 462;
 Best Local Similarity 100.0%; Pred. No. 55;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 MYPT 6
 DB 229 MYPT 233
 RESULT 13
 MM12_RAT STANDARD; PRT; 465 AA.
 ID MM12_RAT
 AC Q63341;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE MACROPHAGE METALLOELASTASE PRECURSOR (EC 3.4.24.65) (MME) (MATRIX
 METALLOPROTEINASE-12) (MMP-12).
 GN MMP12 OR MMEL
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY;
 RA Cossins J., Clements J., Catlin G.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: MAY BE INVOLVED IN TISSUE INJURY AND REMODELING. HAS
 CC SIGNIFICANT ELASTOLYTIC ACTIVITY (BY SIMILARITY).
 CC -!- CATALYTIC ACTIVITY: HYDROLYSIS OF SOLUBLE AND INSOLUBLE ELASTIN.
 CC SPECIFIC CLEAVAGES ARE ALSO PRODUCED AT 14-ALA-|-LEU-15 AND 16-
 CC TYR-|-LEU-17 IN THE B CHAIN OF INSULIN.
 CC -!- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY (BY SIMILARITY).
 CC -!- SIMILARITY: CONTAINS 1 HEMOPEXIN-LIKE DOMAIN.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC
 CC METALLOPROTEASE) ALSO KNOWN AS MATRININ SUBFAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X98517; CAA67142.1; -
 CC HSP; P03956; ICGL.
 CC MEROPS; M10.009; -
 CC InterPro; IPR000130; -
 CC InterPro; IPR000585; -
 CC InterPro; IPR001818; -
 CC Pfam; PF000413; Peptidase_M10; 1.
 CC Pfam; PF00045; hemopexin; 4.
 CC PRINTS; PR00138; MATRININ.
 CC PROSITE; PS00024; HEMOPEXIN; 1.
 CC PROSITE; PS00142; ZINC_PROTEASE; 1.
 CC Hydrolase; Metalloprotease; Glycoprotein; zinc; zymogen; Calcium;
 KW Extracellular matrix; Signal.
 FT SIGNAL 1 21 PROBABLE.
 FT PROPP 22 101 ACTIVATION PEPTIDE (BY SIMILARITY).

```

FT CHAIN 102 465 MACROPHAGE METALLOELASTASE.
FT DOMAIN 275 465 HEMOPEXIN-LIKE.
FT SITE 88 88 CYSTEINE SWITCH (BY SIMILARITY).
FT METAL 214 214 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 215 215 BY SIMILARITY.
FT METAL 218 218 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 224 224 ZINC (CATALYTIC) (BY SIMILARITY).
FT CARBOHYD 313 313 N-GLYCOSYLATED (GLYCAN. . .) (POTENTIAL).
FT DISULFID 278 465 BY SIMILARITY.
SQ SEQUENCE 465 AA; 53738 MW; E779B6014EC6FF68 CRC64;

Query Match 75.6%; Score 31; DB 1; Length 465;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
DB 232 MYPTY 236

RESULT 14
MM18_XENLA
ID MM18_XENLA STANDARD; PRT; 467 AA.
AC O13065;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DE MATRIX METALLOPROTEINASE-18 PRECURSOR (EC 3.4.24.-) (MMP-18)
DE (COLLAGENASE-4) (COLLAGENASE 4) (XCOLA).
GN MMP18 OR COL4.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RA 11;
RP SEQUENCE FROM N.A.
RC TISSUE=Intestine;
RC MEDLINE=97053976; PubMed=8898355;
RA Stoolow M.A., Bauzon D.D., Li J., Sedgwick T., Liang V.C.-T.,
RA "Idg, C.A., Shi Y.-B.;
RT laevis, possible roles during frog development.";
RL Mol. Biol. Cell 7:1471-1483(1996).
CC -1- FUNCTION: CLEAVES COLLAGEN TYPE I. MAY PLAY A ROLE IN LARVAL
CC METAMORPHOSIS. MAY BE INVOLVED IN TAIL RESORPTION.
CC -1- ENZYME REGULATION: REQUIRES CALCIUM AND ZINC FOR ACTIVITY (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: UPREGULATED IN THE TAIL BY THYROID HORMONE.
CC MATRIX (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: EXPRESSED ONLY TRANSIENTLY IN WHOLE ANIMAL, AT
CC TIME WHEN TADPOLE FEEDING BEGINS.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED AT HIGH LEVEL AS THE TADPOLE TAIL
CC RESORBS AND DURING HINDLIMB MORPHOGENESIS AND INTESTINAL
CC REMODELING.
CC -1- SIMILARITY: CONTAINS 1 HEMOPEXIN-LIKE DOMAIN.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC
CC METALLOPROTEASE) ALSO KNOWN AS MATRIXIN SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L76275; AAB53148.1;
CC HSP; P03956; ICGL.
CC MEROPS; M10.018;
CC InterPro; IPR000130;

```

```

DR InterPro; IPR000585;
DR InterPro; IPR001818;
DR Pfam; PF00045; hemopexin; 4.
DR Pfam; PF00413; peptidase_M10; 1.
DR PRINTS; PRO0138; MATRIXIN.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
DR PROSITE; PS00546; CYSTEINE_SWITCH; FALSE_NEG.
DR PROSITE; PS00024; HEMOPEXIN; FALSE_NEG.
KW hydrolase; Metalloprotease; Zinc; Calcium; Zymogen; Signal;
KW signal degradation; Extracellular matrix.
FT SIGNAL 1 17 POTENTIAL.
FT PROPEP 18 99 BY SIMILARITY.
FT CHAIN 100 467 MATRIX METALLOPROTEINASE-18.
FT DOMAIN 277 467 HEMOPEXIN-LIKE.
FT SITE 92 92 CYSTEINE SWITCH (POTENTIAL).
FT METAL 218 218 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 219 219 BY SIMILARITY.
FT METAL 222 222 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 228 228 ZINC (CATALYTIC) (BY SIMILARITY).
FT DISULFID 280 467 POTENTIAL.
SQ SEQUENCE 467 AA; 52812 MW; 4623F6CE3454051 CRC64;

Query Match 75.6%; Score 31; DB 1; Length 467;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
DB 236 MYPTY 240

RESULT 15
MM20_MOUSE
ID MM20_MOUSE STANDARD; PRT; 482 AA.
AC P57748;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DE MATRIX METALLOPROTEINASE-20 PRECURSOR (EC 3.4.24.-) (MMP-20) (ENAMEL
DE METALLOPROTEINASE) (ENAMELYSIN).
GN MMP20.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RA 11;
RP SEQUENCE FROM N.A.
RC STRAIN=129/SVJ;
RC MEDLINE=20079167; PubMed=10610728;
RA Caterina J., Shi J., Krakora S., Bartlett J.D., Engler J.A.,
RA Kozak C.A., Birkedal-Hansen H.;
RT "Isolation, characterization, and chromosomal location of the mouse
RT enamelysin gene.";
RL Genomics 62:308-311(1999).
CC -1- FUNCTION: DEGRADATES AMELOGENIN, THE MAJOR PROTEIN COMPONENT OF THE
CC ENAMEL MATRIX AND TWO OF THE MACROMOLECULES CHARACTERISING THE
CC CARTILAGE EXTRACELLULAR MATRIX: AGGRECAN AND THE CARTILAGE
CC OLIGOMERIC MATRIX PROTEIN (COMP). MAY PLAY A CENTRAL ROLE IN TOOTH
CC ENAMEL FORMATION (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: CLEAVES AGGRECAN AT THE 360-ASN-1-PHE-361 SITE
CC (BY SIMILARITY).
CC -1- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: SECRETED. ASSOCIATED WITH THE EXTRACELLULAR
CC MATRIX (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: EXPRESSED SPECIFICALLY IN THE ENAMEL ORGAN.
CC -1- PTD KUPACTIVATES AT LEAST AT THE 106-ASN-1-TYR-107 SITE (BY
CC SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 HEMOPEXIN-LIKE DOMAIN.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC
CC METALLOPROTEASE) ALSO KNOWN AS MATRIXIN SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; AF156956; AAF28472.1; -
DR EMBL; AF156947; AAF28472.1; JOINED.
DR EMBL; AF156948; AAF28472.1; JOINED.
DR EMBL; AF156949; AAF28472.1; JOINED.
DR EMBL; AF156950; AAF28472.1; JOINED.
DR EMBL; AF156951; AAF28472.1; JOINED.
DR EMBL; AF156952; AAF28472.1; JOINED.
DR EMBL; AF156953; AAF28472.1; JOINED.
DR EMBL; AF156954; AAF28472.1; JOINED.
DR EMBL; AF156955; AAF28472.1; JOINED.
DR EMBL; AF155933; AAF28470.1; -
DR MEROPS; M10.019; -
DR MGD; MGI:1353466; Mmp20.
DR PROSITE; PS00546; CYSTEINE_SWITCH; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
DR PROSITE; PS00024; HEMOPEXIN; FALSE_NEG.
KW Hydrolase; Metalloprotease; Zinc; Calcium; Zymogen; Signal;
KW Extracellular matrix.
FT SIGNAL 1 21 POTENTIAL.
FT PROPEP 22 106 BY SIMILARITY.
FT CHAIN 107 482 MATRIX METALLOPROTEINASE-20.
FT DOMAIN 292 482 HEMOPEXIN-LIKE.
FT SITE 99 99 CYSTEINE SWITCH (POTENTIAL).
FT METAL 225 225 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 226 226 BY SIMILARITY.
FT METAL 229 229 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 235 235 ZINC (CATALYTIC) (BY SIMILARITY).
FT DISULFID 295 482 BY SIMILARITY.
SQ SEQUENCE 482 AA; 54373 MW; 366149FAF2BDC8BB CRC64;

Query Match 75.6%; Score 31; DB 1; Length 482;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
|
|
|
|
Db 243 MYPTY 247

Search completed: September 4, 2001, 15:54:09
Time: 306 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2001, 15:53:35 ; Search time 80.44 Seconds
(without alignments)
13.158 Million cell updates/sec

Title: US-09-630-345-3
Perfect score: 41
Sequence: 1 MYPTYLK 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL16:
1: sp-archaea:
2: sp-bacteria:
3: sp-fungi:
4: sp-human:
5: sp-invertebrate:
6: sp-mammal:
7: sp-mhc:
8: sp-organelle:
9: sp-phage:
10: sp-plant:
11: sp-rodent:
12: sp-unclassified:
13: sp-vertebrate:
14: sp-virus:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35	85.4	1088	4 Q9H2T7	Q9H2T7 homo sapien
2	33	80.5	269	1 Q27159	Q27159 methanobact
3	33	80.5	1069	10 Q9LTZ8	Q9LTZ8 arabidopsis
4	32	78.0	252	2 Q9L875	Q9L875 haemophilus
5	32	78.0	553	11 Q70577	Q70577 mus musculu
6	32	78.0	553	11 Q9R0W2	Q9R0W2 rattus norv
7	32	78.0	593	11 P97558	P97558 rattus norv
8	32	78.0	593	11 P70485	P70485 rattus norv
9	32	78.0	986	4 Q94858	Q94858 homo sapien
10	32	78.0	1346	4 Q9UIZ3	Q9UIZ3 homo sapien
11	31	75.6	40	14 Q9IWZ8	Q9IWZ8 human coxa
12	31	75.6	40	14 Q9IWZ7	Q9IWZ7 human coxa
13	31	75.6	50	14 Q66796	Q66796 human echov
14	31	75.6	53	6 Q9XS83	Q9XS83 equus cabal
15	31	75.6	198	14 Q9Q907	Q9Q907 shope fibro
16	31	75.6	261	4 Q9NRE1	Q9NRE1 homo sapien
17	31	75.6	261	4 Q9NRE7	Q9NRE7 homo sapien
18	31	75.6	261	4 Q9GZ52	Q9GZ52 homo sapien
19	31	75.6	262	5 Q19918	Q19918 caenorhabdi

20	31	75.6	267	6 Q9TV55	Q9TV55 sus scrofa
21	31	75.6	291	5 Q94467	Q94467 dictyosteli
22	31	75.6	307	5 Q9VJ65	Q9VJ65 drosophila
23	31	75.6	332	2 Q51888	Q51888 prevotella
24	31	75.6	378	10 Q9SHH5	Q9SHH5 arabidopsis
25	31	75.6	385	10 Q9SKA2	Q9SKA2 arabidopsis
26	31	75.6	420	4 Q9HAD2	Q9HAD2 homo sapien
27	31	75.6	445	5 Q9VPA2	Q9VPA2 drosophila
28	31	75.6	483	13 Q98857	Q98857 cynops pyrr
29	31	75.6	497	14 Q96640	Q96640 bovine papi
30	31	75.6	502	5 Q18426	Q18426 caenorhabdi
31	31	75.6	516	5 Q18198	Q18198 caenorhabdi
32	31	75.6	565	1 Q29350	Q29350 archaeglob
33	31	75.6	585	5 Q9TXN1	Q9TXN1 caenorhabdi
34	31	75.6	769	11 Q9ESN3	Q9ESN3 mus musculu
35	31	75.6	2185	14 Q88445	Q88445 swine vesic
36	31	73.2	137	14 Q97919	Q97919 human coxa
37	30	73.2	141	5 Q02173	Q02173 caenorhabdi
38	30	73.2	152	2 Q91401	Q91401 pseudomonas
39	30	73.2	155	14 Q98534	Q98534 paramecium
40	30	73.2	158	5 Q9VU21	Q9VU21 drosophila
41	30	73.2	188	5 Q9V5Y8	Q9V5Y8 drosophila
42	30	73.2	238	14 Q9QB94	Q9QB94 yaba monkey
43	30	73.2	248	2 Q48871	Q48871 lactobacill
44	30	73.2	255	14 Q9DHW6	Q9DHW6 yaba-like d
45	30	73.2	262	2 Q9PIA0	Q9PIA0 campylobact

ALIGNMENTS

RESULT 1
Q9H2T7 PRELIMINARY; PRT; 1088 AA.
ID Q9H2T7
AC Q9H2T7
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE RANBP17.

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20568271; PubMed=11024021;
RA Kutay U., Hartmann E., Treichel N., Calado A., Carmo-Fonseca M.,
RA Prehn S., Kraft R., Gorlich D., Bischoff F.R.;
RT "Identification of Two Novel RANGFP-binding Proteins Belonging to the
RT Importin beta Superfamily";
RL J. Biol. Chem. 275:40163-40168(2000).
DR EMBL; AF222747; AAG44255.1;
SQ SEQUENCE 1088 AA; 124374 MW; A95D5599388AEFB CRC64;

Query Match 85.4%; Score 35; DB 4; Length 1088;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYL 7
DB 744 MYPTYL 749

RESULT 2
Q27159 PRELIMINARY; PRT; 269 AA.
ID Q27159
AC Q27159
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE HYPOTHETICAL 31.5 KDA PROTEIN.
GN MTH1087.

```

OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
OC Methanothermobacter.
OX NCBI_TaxID=145262;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DELTA H;
RX MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucette-Stamm L.A., DeLoughery C., Lee H.-M., DuBois J.,
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McDougall S., Shiner G., Goyal A., Pietrowski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Nolling J., Reeve J.N.;
RA "Complete genome sequence of Methanobacterium thermoautotrophicum
RT deltar: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7139-7155(1997).
DR EMBL; AE000879; AB85576.1; -;
DR Hypothetical protein.
SQ SEQUENCE 269 AA; 31513 MW; 4B7D64EBDC6B8EA7 CRC64;

Query Match 80.5%; Score 33; DB 1; Length 269;
Best Local Similarity 83.3%; Pred. No. 52;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 YPTYLK 8
|||||
Db 175 YPTVIK 180

RESULT 3
Q9LTZ8 PRELIMINARY; PRT; 1069 AA.
AC Q9LTZ8
DT 01-OCT-2000 (TRENBLREL. 15, Created)
DT 01-OCT-2000 (TRENBLREL. 15, Last sequence update)
DT 01-MAR-2001 (TRENBLREL. 16, Last annotation update)
DE SIMILARITY TO NEGATIVE REGULATOR OF VESICLE FORMATION.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;
RX Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=20277480; PubMed=10819329;
RA Nakamura Y.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
RT features of the regions of 4,504,864 bp covered by sixty pl and TAC
RT clones."
RL DNA Res. 7:131-135(2000).
DR EMBL; AB024028; BAA95706.2; -;
DR EMBL; AP000381; BAA95706.2; JOINED.
DR InterPro; IPR000379; -;
DR InterPro; IPR000734; -;
DR PROSITE; PS00120; LIPASE_SER; UNKNOWN_1.
SQ SEQUENCE 1069 AA; 118974 MW; 0BA3545568EC6B44 CRC64;

Query Match 80.5%; Score 33; DB 10; Length 1069;
Best Local Similarity 83.3%; Pred. No. 2.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 MYPTYL 7
|||||

```

```

Db 71 MYPTYI 76

RESULT 4
Q9L875 PRELIMINARY; PRT; 252 AA.
AC Q9L875
DT 01-OCT-2000 (TRENBLREL. 15, Created)
DT 01-OCT-2000 (TRENBLREL. 15, Last sequence update)
DT 01-MAR-2001 (TRENBLREL. 16, Last annotation update)
DE BETA 1-4 GLUCOSYLTRANSFERASE LGTF.
GN LGTF
OS Haemophilus ducreyi.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=730;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=35000;
RX MEDLINE=20278119; PubMed=10816485;
RA Filiatrault M.J., Gibson B.W., Schilling B., Sun S., Munson R.S. Jr.,
RA Campagnari A.A.;
RT "Construction and Characterization of Haemophilus ducreyi
RT Lipooligosaccharide (LOS) Mutants Defective in Expression of
RT Heptosyltransferase III and beta1,4-Glucosyltransferase:
RT Identification of LOS Glycoforms Containing Lactosamine Repeats."
RL Infect. Immun. 68:3352-3361(2000).
DR EMBL; AF215936; AAF72876.1; -;
DR InterPro; IPR001173; -;
DR Pfam; PF00535; Glycos_transf_2; 1.
KW Transferase.
SQ SEQUENCE 252 AA; 28736 MW; 20297CB79C070AD4 CRC64;

Query Match 78.0%; Score 32; DB 2; Length 252;
Best Local Similarity 83.3%; Pred. No. 78;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 MYPTYL 7
|||||
Db 134 LYPTYL 139

RESULT 5
Q70577 PRELIMINARY; PRT; 553 AA.
AC Q70577
DT 01-AUG-1998 (TRENBLREL. 07, Created)
DT 01-AUG-1998 (TRENBLREL. 07, Last sequence update)
DT 01-MAR-2001 (TRENBLREL. 16, Last annotation update)
DE ORGANIC CATION TRANSPORTER 2.
GN SLC22A2 OR OCT2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129; TISSUE=LIVER;
RX Mooslehner K.A., Allen N.D.;
RX Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
CC -|- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -|- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
DR EMBL; AJ006036; CAA06827.1; -;
DR MGD; MGI:1335072; SLC22a2.
DR InterPro; IPR001066; -;
DR Pfam; PF00083; sugar_tr; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_2.
KW Transmembrane.
SQ SEQUENCE 553 AA; 61830 MW; 1E9744F0D5415483 CRC64;

Query Match 78.0%; Score 32; DB 11; Length 553;

```


Best Local Similarity 57.1%; Pred. No. 1.8e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 4; Conservative 3;

QY 2 MYPTYLK 8
:||||:
Db 457 LYPTIIR 463

RESULT 6
Q9ROW2 PRELIMINARY; PRT; 555 AA.
AC Q9ROW2
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE ORGANIC CATION TRANSPORTER OCT2R.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
[1]
SEQUENCE FROM N.A.
RX TISSUE=KIDNEY;
RC MEDLINE=99316020; PubMed=10385678;
RA Gruendemann D., Lieblich G., Kiefer N., Koester S., Schoemig E.;
RT "Selective substrates for non-neuronal monoamine transporters."; Mol. Pharmacol. 56:1-10(1999).
DR EMBL; Y13154; CAB52215.1; -;
DR InterPro: IPR001066; -;
DR Pfam: PF00083; sugar_tr; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_2.
SQ SEQUENCE 555 AA; 62342 MW; 29521969AEIAC206 CRC64;

Query Match 78.0%; Score 32; DB 11; Length 555;
Best Local Similarity 57.1%; Pred. No. 1.8e+02;

Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
:||||:
Db 457 LYPTIIR 463

RESULT 7
P97558 PRELIMINARY; PRT; 593 AA.
AC P97558
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE ORGANIC CATION TRANSPORTER.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
[1]
SEQUENCE FROM N.A.
RP TISSUE=KIDNEY;
RA Gorboulev V.G., Koepsell H.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
CC -|- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -|- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
DR EMBL; X98334; CAA66979.1; -;
DR InterPro: IPR001066; -;
DR Pfam: PF00083; sugar_tr; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_2.
KW Ion transport; Transmembrane.
SQ SEQUENCE 593 AA; 66100 MW; 36C3E1B5DC057790 CRC64;

Query Match 78.0%; Score 32; DB 11; Length 593;
Best Local Similarity 57.1%; Pred. No. 2e+02;

Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
:||||:
Db 457 LYPTIIR 463

RESULT 8
P70485 PRELIMINARY; PRT; 593 AA.
AC P70485
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE ORGANIC CATION TRANSPORTER OCT2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
[1]
SEQUENCE FROM N.A.
RP STRAIN=SPRAGUE-DAWLEY; TISSUE=KIDNEY;
RC MEDLINE=96295517; PubMed=8702418;
RX Okuda M., Saito H., Urakami Y., Takano M., Inui K.;
RA "cDNA cloning and functional expression of a novel rat kidney organic cation transporter, OCT2."; Biochem. Biophys. Res. Commun. 224:500-507(1996).
CC -|- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -|- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
DR EMBL; D83044; BAA11754.1; -;
DR InterPro: IPR001066; -;
DR Pfam: PF00083; sugar_tr; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_2.
KW Transmembrane.
SQ SEQUENCE 593 AA; 66080 MW; 36C1044E0C04B443 CRC64;

Query Match 78.0%; Score 32; DB 11; Length 593;
Best Local Similarity 57.1%; Pred. No. 2e+02;

Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
:||||:
Db 457 LYPTIIR 463

RESULT 9
O94858 PRELIMINARY; PRT; 986 AA.
AC O94858
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE KIAA0758 PROTEIN (FRAGMENT).
GN KIAA0758.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
RP TISSUE=BRAIN;
RC MEDLINE=99087487; PubMed=9872452;
RX Nagase T., Ishikawa K., Suyama M., Kikuno R., Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XI. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro."; DNA Res. 5:277-286(1998).
DR EMBL; AB018301; BAA34478.1; -;
DR InterPro: IPR000203; -;
DR Pfam: PF00083; -;
DR InterPro: IPR000832; -;
DR Pfam: PF00002; 7tm_2; 1.
DR Pfam: PF01825; GPS; 1.

DR PRINTS; PR00249; GPCRSECRETIN.
 DR SMART; SM00303; GPS; 1.
 FT NON_TER 1
 SQ SEQUENCE 986 AA; 108720 MW; EF4875B03BAA0E7D CRC64;

Query Match 78.0%; Score 32; DB 4; Length 986;
 Best Local Similarity 85.7%; Pred. No. 3.4e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
 Db 384 MLPTYLK 390

RESULT 10

Q9UIZ3
 ID Q9UI23 PRELIMINARY; PRT; 1346 AA.
 AC Q9UI23
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DE DJ365012.1 (KIAA0758 PROTEIN).
 GN DJ365012.1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN (1)_TaxID=9606;
 RP SEQUENCE FROM N.A.
 RA Seba H.
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX DOMAIN
 DR EMBL; ALA96772; CAB61578.1; -.
 DR InterPro: IPR000082; -.
 DR InterPro: IPR000083; -.
 DR InterPro: IPR000082; -.
 DR InterPro: IPR000082; -.
 DR InterPro: IPR000082; -.
 DR Pfam: PF00002; 7tm2; 1.
 DR Pfam: PF00047; 4tm2; 1.
 DR Pfam: PF01825; GPS; 1.
 DR PRINTS; PR00249; GPCRSECRETIN.
 DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
 DR SMART; SM00200; SEA; 1.
 SQ SEQUENCE 1346 AA; 149456 MW; 72A9D02B08218A60 CRC64;

Query Match 78.0%; Score 32; DB 4; Length 1346;
 Best Local Similarity 85.7%; Pred. No. 4.7e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
 Db 744 MLPTYLK 750

RESULT 11

Q9IWZ8
 ID Q9IW28 PRELIMINARY; PRT; 40 AA.
 AC Q9IW28
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE VP2 PROTEIN (FRAGMENT).
 GN VP2.
 OS Human coxsackievirus B5.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Enterovirus.
 OX NCBI_TaxID=12074;
 RN (1)
 RP SEQUENCE FROM N.A.

RC STRAIN-R8597;
 RA Casas I., Palacios G.F., Trallero G., Cisterna D., Freire M.C.,
 RA "Memorio A.;
 RT "Molecular characterization of human enteroviruses in clinical samples
 RT by three different RT nested PCR assays and direct sequencing of
 RT amplified products.";
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF252193; AAF71330.1; -.
 DR InterPro: IPR001676; -.
 DR Pfam: PF00073; rhv; 1.
 DR Non_TER 1
 FT NON_TER 40
 SQ SEQUENCE 40 AA; 4388 MW; 7F998A9865EF9F67 CRC64;

Query Match 75.6%; Score 31; DB 14; Length 40;
 Best Local Similarity 71.4%; Pred. No. 17;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
 Db 26 VMPYTLK 32

RESULT 12

Q9IWZ7
 ID Q9IWZ7 PRELIMINARY; PRT; 40 AA.
 AC Q9IWZ7
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE VP2 PROTEIN (FRAGMENT).
 GN VP2.
 OS Human coxsackievirus B5.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Enterovirus.
 OX NCBI_TaxID=12074;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-R8497;
 RA Casas I., Palacios G.F., Trallero G., Cisterna D., Freire M.C.,
 RA "Memorio A.;
 RT "Molecular characterization of human enteroviruses in clinical samples
 RT by three different RT nested PCR assays and direct sequencing of
 RT amplified products.";
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF252194; AAF71331.1; -.
 DR InterPro: IPR001676; -.
 DR Pfam: PF00073; rhv; 1.
 DR Non_TER 1
 FT NON_TER 40
 SQ SEQUENCE 40 AA; 4388 MW; 7F998A9865EF9F67 CRC64;

Query Match 75.6%; Score 31; DB 14; Length 40;
 Best Local Similarity 71.4%; Pred. No. 17;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
 Db 26 VMPYTLK 32

RESULT 13

Q66796
 ID Q66796 PRELIMINARY; PRT; 50 AA.
 AC Q66796
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE CAPSID PROTEIN VP2 (FRAGMENT).
 GN VP2.
 OS Human echovirus 20.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Enterovirus.

OC NCBI_TaxID=47508;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Jv-1;
 RX MEDLINE=96203959; PubMed=8627260;
 RA Huttunen P., Santti J., Pulli T., Hyypia T.;
 RT "The major echovirus group is genetically coherent and related to
 coxsackie B viruses.";
 RL J. Gen. Virol. 77:715-725(1996).
 DR EMBL: X89549; CAA61727.1; -;
 DR InterPro: IPR001676; -;
 DR Pfam: PF00073; rhv; 1.
 FT NON_TER 1
 FT NON_TER 50
 FT NON_TER 50
 SQ SEQUENCE 50 AA; 5448 MW; BBFC58B2511FC2E7 CRC64;

Query Match 75.6%; Score 31; DB 14; Length 50;
 Best Local Similarity 71.4%; Pred. No. 22;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
 DB 31 VWPTYLK 37

RESULT 14

QX9S83 PRELIMINARY; PRT; 53 AA.
 AC QX9S83;
 DT 01-NOV-1999 (TReMBLrel. 12, Created)
 DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
 DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
 DE TUBBY PROTEIN (FRAGMENT).
 GN TUB.
 OS Equus caballus (Horse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
 OX NCBI_TaxID=9796;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99160468; PubMed=10051323;
 RA Caetano A.R., Pomp D., Murray J.D., Bowling A.T.;
 RT "Comparative mapping of 18 equine type I genes assigned by somatic
 cell hybrid analysis.";
 RL Mamm. Genome 10:271-276(1999).
 RI EMBL: AF097580; AAD25983.1; -;
 HS SP: P50586; IC82.
 DR InterPro: IPR000007; -;
 DR Pfam: PF01167; Tub; 1.
 FT NON_TER 1
 FT NON_TER 53
 FT NON_TER 53
 SQ SEQUENCE 53 AA; 6244 MW; 13A9ED8C69500D3E CRC64;

Query Match 75.6%; Score 31; DB 6; Length 53;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTL 6
 DB 16 MYPTL 20

RESULT 15

Q9Q907 PRELIMINARY; PRT; 198 AA.
 ID Q9Q907
 AC Q9Q907;
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)

DE GP064R.
 GN S064R.
 OS Shope fibroma virus (strain Kasza) (SFV).
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
 OC Leporipoxvirus.
 OX NCBI_TaxID=10272;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KASZA;
 RX MEDLINE=84165064; PubMed=6323741;
 RA Delange A.M., Macaulay C., Block W., Mueller T., McFadden G.;
 RT "Tumorigenic poxviruses: construction of the composite physical map of
 the Shope fibroma virus genome.";
 RL J. Virol. 50:408-416(1984).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KASZA;
 RX MEDLINE=92074222; PubMed=1660196;
 RA Strayer D.S., Jerng H.H., O'Connor K.;
 RT "Sequence and analysis of a portion of the genomes of Shope fibroma
 virus and malignant rabbit fibroma virus that is important for viral
 replication in lymphocytes.";
 RL Virology 185:585-595(1991).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KASZA;
 RX MEDLINE=20032074; PubMed=10562495;
 RA Willer D.O., McFadden G., Evans D.H.;
 RT "The complete genome sequence of shope (Rabbit) fibroma virus.";
 RL Virology 264:319-343(1999).
 DR EMBL: AF170722; AAF17946.1; -;
 SQ SEQUENCE 198 AA; 23282 MW; 9D181052C1ED42C5 CRC64;

Query Match 75.6%; Score 31; DB 14; Length 198;
 Best Local Similarity 83.3%; Pred. No. 96;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYL 7
 DB 110 MYPTFL 115

Search completed: September 4, 2001, 15:53:36
 Job time: 289 sec

Tue Sep 4 15:56:21 2001

us-09-630-345-3.rspt

Page 6

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2001, 15:50:25 ; Search time 77.49 Seconds
(without alignments)
6.259 Million cell updates/sec

Title: US-09-630-345-4
Perfect score: 41
Sequence: 1 XMYPTYLK 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: A_Geneseq_0601.*

- 1: /SID88/gcgdata/geneseq/geneseqp/AA1980.DAT.*
- 2: /SID88/gcgdata/geneseq/geneseqp/AA1981.DAT.*
- 3: /SID88/gcgdata/geneseq/geneseqp/AA1982.DAT.*
- 4: /SID88/gcgdata/geneseq/geneseqp/AA1983.DAT.*
- 5: /SID88/gcgdata/geneseq/geneseqp/AA1984.DAT.*
- 6: /SID88/gcgdata/geneseq/geneseqp/AA1985.DAT.*
- 7: /SID88/gcgdata/geneseq/geneseqp/AA1986.DAT.*
- 8: /SID88/gcgdata/geneseq/geneseqp/AA1987.DAT.*
- 9: /SID88/gcgdata/geneseq/geneseqp/AA1988.DAT.*
- 10: /SID88/gcgdata/geneseq/geneseqp/AA1990.DAT.*
- 11: /SID88/gcgdata/geneseq/geneseqp/AA1991.DAT.*
- 12: /SID88/gcgdata/geneseq/geneseqp/AA1992.DAT.*
- 13: /SID88/gcgdata/geneseq/geneseqp/AA1993.DAT.*
- 14: /SID88/gcgdata/geneseq/geneseqp/AA1994.DAT.*
- 15: /SID88/gcgdata/geneseq/geneseqp/AA1995.DAT.*
- 16: /SID88/gcgdata/geneseq/geneseqp/AA1996.DAT.*
- 17: /SID88/gcgdata/geneseq/geneseqp/AA1997.DAT.*
- 18: /SID88/gcgdata/geneseq/geneseqp/AA1998.DAT.*
- 19: /SID88/gcgdata/geneseq/geneseqp/AA1999.DAT.*
- 20: /SID88/gcgdata/geneseq/geneseqp/AA2000.DAT.*
- 21: /SID88/gcgdata/geneseq/geneseqp/AA2001.DAT.*
- 22: /SID88/gcgdata/geneseq/geneseqp/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	97.6	8	AAW50937	Vasoactive intesti
2	40	97.6	8	AAW50937	Vasoactive intesti
3	40	97.6	8	AAW50937	Amino acid sequenc
4	35	85.4	8	AAW50937	Amino acid sequenc
5	35	85.4	8	AAW50937	Amino acid sequenc
6	35	85.4	450	AAW50937	Human secreted pro
7	32	78.0	807	AAW50937	Human GPCR protein
8	32	78.0	986	AAW50937	Human brain-derive
9	31	75.6	8	AAW50937	Mammalian tub prot
10	31	75.6	12	AAW50937	Peptidomimetic cap
11	31	75.6	58	AAW50937	Human 5' EST seque

12	31	75.6	173	22	AAW74618	Human 72 kDa gelat
13	31	75.6	173	22	AAW68682	Matrilysin catalyt
14	31	75.6	261	19	AAW76253	Human matrilysin-1
15	31	75.6	264	19	AAW76254	Human matrilysin p
16	31	75.6	271	17	AAW96211	Human recombinant
17	31	75.6	297	21	AAW43585	Human cancer assoc
18	31	75.6	325	14	AAW42750	ced-4 gene II prod
19	31	75.6	459	18	AAW36485	Mouse TUB Form I.
20	31	75.6	459	21	AAW26400	Mouse tub Form I p
21	31	75.6	460	18	AAW36488	Human TUB Form 6.
22	31	75.6	460	21	AAW26903	Human TUB Form 6 p
23	31	75.6	462	22	AAW49983	Murine macrophage
24	31	75.6	469	22	AAW10655	BPV1 L1 fusion pro
25	31	75.6	469	21	AAW10660	BPV2 L1 fusion pro
26	31	75.6	483	20	AAW93163	Human MMP-20 prote
27	31	75.6	495	6	AAW51099	Sequence of bovine
28	31	75.6	495	20	AAW88482	Bovine papillomavi
29	31	75.6	505	18	AAW36486	Mouse TUB Form II.
30	31	75.6	505	18	AAW10728	Mouse tub gene pro
31	31	75.6	505	19	AAW54367	Mouse tub polypt
32	31	75.6	505	20	AAW75450	Mouse wild type tu
33	31	75.6	505	21	AAW26901	Mouse tub Form II
34	31	75.6	506	18	AAW36497	Human TUB Form 4.
35	31	75.6	506	18	AAW10729	Human tub gene pro
36	31	75.6	506	19	AAW54368	Human tub polypt
37	31	75.6	506	20	AAW75451	Human wild type tu
38	31	75.6	506	21	AAW26910	Human TUB form 4 5
39	31	75.6	512	18	AAW36496	Human TUB Form 3.
40	31	75.6	512	21	AAW26912	Human TUB form 3 5
41	31	75.6	518	18	AAW36495	Human TUB Form 2.
42	31	75.6	518	21	AAW26909	Human TUB form 2 5
43	31	75.6	561	18	AAW36489	Human TUB Form 1.
44	31	75.6	561	21	AAW26904	Human TUB Form 1 p
45	30	73.2	73	14	AAW40041	Filistata peptide

ALIGNMENTS

RESULT 1
AAW50937
ID AAW50937 standard; peptide; 8 AA.
XX
AC AAW50937;
XX
DT 31-JUL-1998 (first entry)
XX
DE Vasoactive intestinal peptide receptor binding inhibitor (VIP2).
XX
KW Vasoactive intestinal peptide; VIP; antagonist; somatostatin; bombesin;
KW Substance P; cancer; inhibition.
XX
OS Synthetic.
XX
PN EP835662-A2.
XX
PD 15-APR-1998.
XX
PF 11-DEC-1996; 96EP-0309012.
XX
PR 08-OCT-1996; 96US-0727679.
PR 16-AUG-1996; 96IN-0001822.
XX
(NAIM-) NAT INST IMMUNOLOGY.
XX
PI Jaggi M, Mukherjee R;
XX
DR WPI; 1998-208959/19.
XX
PT Composition containing analogues of vasoactive intestinal peptide,
PT somatostatin - bombesin and substance P, for treatment of tumours
PT and for inhibiting over-expression of these peptide(s)
XX

Claim 1; Page 4; 49pp; English.

PS The invention relates to a new composition which comprises: (i) the
 CC somatostatin analogue SOM2 AGCKNFRDWTPTSDC (3-14 disulphide bridge),
 CC and (ii) at least 4 of the peptides: antagonist of vasoactive
 CC intestinal peptide (VIP1); VIP receptor-binding inhibitor (VIP2); VIP
 CC receptor antagonist (VIP3); somatostatin analogue (SOM1); bombesin
 CC antagonist (BOM1) and substance P antagonist (SP1). Also claimed are
 CC more general compositions containing peptide analogues of somatostatin,
 CC VIP, bombesin and substance P. The compositions are used in human or
 CC veterinary medicine: (a) to kill (or inhibit multiplication of) tumour
 CC or cancer cells, particularly for treatment of leukaemia, lymphoma,
 CC adenocarcinoma of stomach, pancreas or prostate, or cancer of lung,
 CC breast, kidney or particularly rectum and colon, and (b) to prevent,
 CC inhibit or modulate over-expression of, e.g. VIP. A wide range of cancer
 CC cells express receptors for VIP, somatostatin, bombesin and/or substance
 CC P. The present sequence represents VIP receptor-binding inhibitor
 CC (VIP2).

SQ Sequence 8 AA;

Query Match 97.6%; Score 40; DB 19; Length 8;
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
 Db 2 myptylk 8
 |||||

RESULT 2

ID AAB08298 standard; peptide; 8 AA.

XX AAB08298;

DT 04-DEC-2000 (first entry)

XX Vasoactive intestinal peptide (VIP) analogue VIP2.

XX Vasoactive intestinal peptide; VIP; analogue; somatostatin; SOM1; SOM2;
 KW VIP1; VIP2; VIP3; BOM1; bombesin; SP1; substance P; MuJ-7; tumour growth;
 KW tumour angiogenesis; metastasis; cancer; angiogenesis; adenocarcinoma;
 KW leukaemia; lymphoma.

XX Synthetic.

PN WO200047221-A1.

XX 17-AUG-2000.

PF 11-FEB-2000; 2000WO-US03559.

PR 11-FEB-1999; 99US-0248381.

XX (NAIM-) NAT INST IMMUNOLOGY.

PA (DABU-) DABUR RES FOUND.

PA (CORD/) CORD J I.

XX Mukherjee R, Jaggi M, Prasad S, Burman AC, Rajendran P, Mathur A;
 PI Singh AT;

XX WPT; 2000-549083/50.

XX Novel therapeutically active composition comprising at least 5
 PT peptides, useful for treating angiogenesis especially as a result of
 PT adenocarcinomas -

PS Disclosure; Page 8; 42pp; English.

XX The present sequence represents an analogue of vasoactive intestinal
 CC peptide (VIP). The specification describes therapeutically active

CC compositions comprising at least one analogue of somatostatin (chosen
 CC from SOM1 and SOM2), and at least four analogues chosen from VIP1 (a
 CC VIP antagonist), VIP2 (a VIP receptor binding inhibitor), VIP3 (a VIP
 CC receptor antagonist), BOM1 (a bombesin antagonist), and SP1 (a substance
 CC P antagonist). The combination of these 7 analogues is known as MuJ-7.
 CC MuJ-7 is used as an anticancer drug to restrict tumour growth and spread
 CC by inhibiting tumour angiogenesis. MuJ-7, in addition, inhibits
 CC metastasis through its angiogenic activity in all cancers. The
 CC peptides are useful for the treatment and prevention of angiogenesis,
 CC especially as a result of adenocarcinomas of the colon, breast, lung,
 CC prostate, kidney, leukemias or lymphomas.

SQ Sequence 8 AA;

Query Match 97.6%; Score 40; DB 21; Length 8;
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
 Db 2 myptylk 8
 |||||

RESULT 3

ID AAB08306 standard; peptide; 8 AA.

XX AAB08306;

DT 04-DEC-2000 (first entry)

XX Amino acid sequence of an antiangiogenic peptide.

XX Vasoactive intestinal peptide; VIP; analogue; somatostatin; SOM1; SOM2;
 KW VIP1; VIP2; VIP3; BOM1; bombesin; SP1; substance P; MuJ-7; tumour growth;
 KW tumour angiogenesis; metastasis; cancer; angiogenesis; adenocarcinoma;
 KW leukaemia; lymphoma.

XX Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 1 /note= "Optionally D-form residue"

FT Misc-difference 3 /note= "Optionally D-form residue"

FT Misc-difference 6 /note= "Optionally D-form residue"

FT Misc-difference 7 /note= "Optionally D-form residue"

FT Misc-difference 7 /note= "Optionally D-form residue"

XX WO200047221-A1.

XX 17-AUG-2000.

PF 11-FEB-2000; 2000WO-US03559.

PR 11-FEB-1999; 99US-0248381.

XX (NAIM-) NAT INST IMMUNOLOGY.

PA (DABU-) DABUR RES FOUND.

PA (CORD/) CORD J I.

XX Mukherjee R, Jaggi M, Prasad S, Burman AC, Rajendran P, Mathur A;
 PI Singh AT;

XX WPT; 2000-549083/50.

XX Novel therapeutically active composition comprising at least 5
 PT peptides, useful for treating angiogenesis especially as a result of
 PT adenocarcinomas -

PS Claim 11; Page 31; 42pp; English.

XX AAB08304-15 represent peptides which have an antiangiogenic effect. The
 CC specification describes therapeutically active compositions comprising
 CC at least one analogue of somatostatin (chosen from SOM1 and SOM2), and
 CC at least four analogues chosen from vasoactive intestinal peptide (VIP)
 CC 1 (a VIP antagonist), VIP2 (a VIP receptor binding inhibitor), VIP3 (a
 CC VIP receptor antagonist), BOM1 (a bombesin antagonist), and SP1 (a
 CC substance P antagonist). The combination of these 7 analogues is known as
 CC MuJ-7. MuJ-7 is used as an anticancer drug to restrict tumour growth and
 CC metastasis through its antiangiogenic activity in all cancers. The
 CC peptides are useful for the treatment and prevention of angiogenesis,
 CC especially as a result of adenocarcinomas of the colon, breast, lung,
 CC prostate, kidney, leukemias or lymphomas.
 XX SQ Sequence 8 AA;

Query Match 97.6%; Score 40; DB 21; Length 8;
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
 |||||
 Db 2 myptylk 8

RESULT 4
 AAB08304
 ID AAB08304 standard; peptide; 8 AA.
 XX AC AAB08304;
 XX DT 04-DEC-2000 (first entry)
 XX DE Amino acid sequence of an antiangiogenic peptide.
 XX KW Vasoactive intestinal peptide; VIP; analogue; somatostatin; SOM1; SOM2;
 KW VIP1; VIP2; VIP3; BOM1; bombesin; SP1; substance P; MuJ-7; tumour growth;
 KW tumour angiogenesis; metastasis; cancer; angiogenesis; adenocarcinoma;
 KW leukaemia; lymphoma.
 XX OS Synthetic.
 XX FH Key Location/Qualifiers
 FT Modified-site 1 /label= Aib
 FT /note= "alpha-aminoisobutyric acid"
 FT Modified-site 7 /label= Aib, Deg
 FT /note= "alpha-aminoisobutyric acid or alpha, alpha
 FT diethyl glycine"
 XX PN WO200047221-A1.
 XX PD 17-AUG-2000.
 XX PF 11-FEB-2000; 2000WO-US03559.
 XX PR 11-FEB-1999; 99US-0248381.
 XX (NAIM-) NAT INST IMMUNOLOGY.
 PA (DABU-) DABUR RES FOUND.
 PA (CORD/) CORD J I.
 XX Mukherjee R, Jaggi M, Prasad S, Burman AC, Rajendran P, Mathur A;
 PI Singh AT;
 XX WPI; 2000-549083/50.
 XX Novel therapeutically active composition comprising at least 5
 PT peptides, useful for treating angiogenesis especially as a result of
 PT adenocarcinomas -

XX Claim 18; Page 36; 42pp; English.
 XX AAB08304-15 represent peptides which have an antiangiogenic effect. The
 CC specification describes therapeutically active compositions comprising
 CC at least one analogue of somatostatin (chosen from SOM1 and SOM2), and
 CC at least four analogues chosen from vasoactive intestinal peptide (VIP)
 CC 1 (a VIP antagonist), VIP2 (a VIP receptor binding inhibitor), VIP3 (a
 CC VIP receptor antagonist), BOM1 (a bombesin antagonist), and SP1 (a
 CC substance P antagonist). The combination of these 7 analogues is known as
 CC MuJ-7. MuJ-7 is used as an anticancer drug to restrict tumour growth and
 CC metastasis through its antiangiogenic activity in all cancers. The
 CC peptides are useful for the treatment and prevention of angiogenesis,
 CC especially as a result of adenocarcinomas of the colon, breast, lung,
 CC prostate, kidney, leukemias or lymphomas.
 XX SQ Sequence 8 AA;

Query Match 85.4%; Score 35; DB 21; Length 8;
 Best Local Similarity 85.7%; Pred. No. 3.4e+05;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
 |||||
 Db 2 myptylk 8

RESULT 5
 AAB08305
 ID AAB08305 standard; peptide; 8 AA.
 XX AC AAB08305;
 XX DT 04-DEC-2000 (first entry)
 XX DE Amino acid sequence of an antiangiogenic peptide.
 XX KW Vasoactive intestinal peptide; VIP; analogue; somatostatin; SOM1; SOM2;
 KW VIP1; VIP2; VIP3; BOM1; bombesin; SP1; substance P; MuJ-7; tumour growth;
 KW tumour angiogenesis; metastasis; cancer; angiogenesis; adenocarcinoma;
 KW leukaemia; lymphoma.
 XX OS Synthetic.
 XX FH Key Location/Qualifiers
 FT Misc-difference 1 /note= "D-form residue"
 FT Modified-site 7 /label= Aib, Deg
 FT /note= "alpha-aminoisobutyric acid or alpha, alpha
 FT diethyl glycine"
 XX PN WO200047221-A1.
 XX PD 17-AUG-2000.
 XX PF 11-FEB-2000; 2000WO-US03559.
 XX PR 11-FEB-1999; 99US-0248381.
 XX (NAIM-) NAT INST IMMUNOLOGY.
 PA (DABU-) DABUR RES FOUND.
 PA (CORD/) CORD J I.
 XX Mukherjee R, Jaggi M, Prasad S, Burman AC, Rajendran P, Mathur A;
 PI Singh AT;
 XX WPI; 2000-549083/50.
 XX Novel therapeutically active composition comprising at least 5
 PT peptides, useful for treating angiogenesis especially as a result of

adenocarcinomas -

Claim 11; Page 31; 42pp; English.

AA08304-15 represent peptides which have an antiangiogenic effect. The specification describes therapeutically active compositions comprising at least one analogue of somatostatin (chosen from SOM1 and SOM2), and at least four analogues chosen from vasoactive intestinal peptide (VIP), VIP1 (a VIP antagonist), VIP2 (a VIP receptor binding inhibitor), VIP3 (a VIP receptor antagonist), SOM1 (a bombesin antagonist), and SPl (a substance P antagonist). The combination of these 7 analogues is known as MuJ-7. MuJ-7 is used as an anticancer drug to restrict tumour growth and spread by inhibiting tumour angiogenesis. MuJ-7, in addition, inhibits metastasis through its antiangiogenic activity in all cancers. The peptides are useful for the treatment and prevention of angiogenesis, especially as a result of adenocarcinomas of the colon, breast, lung, prostate, kidney, leukemias or lymphomas.

Sequence 8 AA:

Query Match 85.4%; Score 35; DB 21; Length 8;
Best Local Similarity 85.7%; Pred. NO. 3.4e+05;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
| | | | |
Db 2 myptyxk 8

RESULT 6
AA063104
ID AA063104 standard; Protein; 450 AA.
XX
AC AA063104;
XX
DT 26-MAR-2001 (first entry)
XX
DE Human secreted protein sequence encoded by gene 18 SEQ ID NO:114.
XX
KW Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;
KW antirheumatic; antiproliferative; cytostatic; cardiant; vasotrophic;
KW cerebroprotective; nootropic; neuroprotective; antibacterial; virucide;
KW fungicide; ophthalmological; vulnary; gene therapy; neoplasm;
KW autoimmune disease; rheumatoid arthritis; hyperproliferative disorder;
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW cerebral ischaemia; angiogenesis; nervous system disorder; infection;
KW Alzheimer's disease; ocular disorder; corneal infection; wound healing;
KW skin aging; food additive; preservative.
XX
OS Homo sapiens.
XX
PN WO200061748-A1.
XX
PD 19-OCT-2000.
XX
PF 06-APR-2000; 2000WO-US08982.
XX
PR 09-APR-1999; 99US-0128696.
PR 14-JAN-2000; 2000US-0176069.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM, Komatsoulis G;
XX
PI WPI; 2000-638566/61.
XX
XX New nucleic acid molecules encoding 48 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions and
PT used as food additives or preservatives -
XX
PS Disclosure; Page 453-455; 480pp; English.
XX

AA022316 to AA022363 encode the human secreted proteins given in AAB63049 to AAB63096. AAB63097 to AAB63132 represent more human secreted proteins and polypeptides homologous to them. Human secreted proteins have activities based on the tissues and cells the genes are expressed in. Examples of activities include: immunosuppressive; antiarthritic; antirheumatic; antiproliferative; cytostatic; cardiant; vasotrophic; cerebroprotective; nootropic; neuroprotective; antibacterial; virucide; fungicide; ophthalmological; and vulnary. The polynucleotides and proteins can be used to prevent, treat or ameliorate a medical condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used in diagnosing a pathological condition or susceptibility to a pathological condition. Disorders which are diagnosed or treated include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g. Alzheimer's disease, infections caused by bacteria, viruses and fungi and ocular disorders e.g. corneal infection. The polypeptides can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, and in supporting cell culture of primary tissues, to regenerate tissues and in chemotaxis. The polypeptides can also be used as a food additive or preservative to increase or decrease storage capabilities, fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors and other nutritional components. AA022307 to AA022315 and AAB63048 represent sequences used in the exemplification of the present invention.

Sequence 450 AA;

Query Match 85.4%; Score 35; DB 21; Length 450;
Best Local Similarity 100.0%; Pred. NO. 40;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYL 7
| | | | |
Db 106 myptyl 111

RESULT 7
AA057288
ID AA057288 standard; Protein; 807 AA.
XX
AC AA057288;
XX
DT 05-JUN-2000 (first entry)
XX
DE Human GPCR protein (HGPRP) sequence (clone ID 3036563).
XX
KW Human; G protein coupled protein receptor; HGPRP; cell proliferation;
KW neurological; immune disorder; cytostatic; anti-arteriosclerotic;
KW anti-atherosclerotic; hepatotropic; antiinflammatory; virucide; leukemia;
KW immunomodulatory; anemia; asthma; gastrointestinal; anti-epileptic;
KW anti-Alzheimer's; anti-Parkinsonian; gene therapy.
XX
OS Homo sapiens.
XX
PN WO200015793-A2.
XX
PD 23-MAR-2000.
XX
PF 17-SEP-1999; 99WO-US20958.
XX
PR 17-SEP-1998; 98US-0156513.
XX
XX (INCY-) INCYTE PHARM INC.
XX
XX Bandman O, Lal P, Tang YT, Corley NC, Guegler KJ, Gorgone GA;
PI Baughn MR;
XX
XX WPI; 2000-271432/23.
XX
XX N-PSDB; AA290526.
XX

PT Human G protein coupled protein receptor peptides useful for the
 PT prevention, diagnosis and treatment of cell proliferative, neurological
 PT and immune disorders -

PS Claim 1; Page 65-67; 71pp; English.

XX The invention provides human G protein coupled protein receptor (HGPRP)
 CC polypeptides and polynucleotides encoding them. The polypeptides can be
 CC produced by standard recombinant methodology. The polynucleotides and
 CC polypeptides may be used in the prevention, treatment and diagnosis of
 CC diseases associated with their inappropriate expression. Diseases that
 CC can be treated are cell proliferative disorders (e.g. arteriosclerosis,
 CC atherosclerosis and hepatitis), cancers (e.g. leukemia, melanomas and
 CC adenocarcinoma), immune disorders (e.g. anemia, asthma and Crohn's
 CC disease) and neurological disorders (e.g. epilepsy, Alzheimer's disease
 CC and Parkinson's disease). The anti-HGPRP antibodies may also be used as
 CC diagnostic agents for detecting the presence of HGPRP polypeptides in
 CC samples (e.g. by enzyme linked immunosorbent assay (ELISA)). Sequences
 CC AAY57283-288 represent the HGPRP polypeptides.

XX SQ Sequence 807 AA;

Query Match 78.0%; Score 32; DB 21; Length 807;
 Best Local Similarity 85.7%; Pred. NO. 2.9e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
 I I I I I
 Db 205 mlptylk 211

RESULT 8
 AAY40440
 ID AAY40440 standard; Protein; 986 AA.

XX AC AAY40440;

XX DT 26-MAY-2000 (first entry)

XX DE Human brain-derived G-protein coupled receptor protein.

XX KW G-protein coupled receptor protein; human; brain; gene therapy;
 KW genetic disease; screening assay.

XX OS Homo sapiens.

XX PN WO200008053-A1.

XX PD 17-FEB-2000.

XX PF 05-AUG-1999; 99WO-JP04233.

XX PR 07-AUG-1998; 98JP-0225059.

XX PA (TAKE) TAKEDA CHEM IND LTD.

XX PA (KAZU-) KAZUSA DNA RES INST.

XX PI Ohara O, Nagase T, Nomura N, Mogi S, Yamamoto K, Kurokawa T;

XX DR WPI; 2000-195555/17.

XX DR N-PSDB; AAZ87930.

XX Human brain-derived G-protein coupled receptor protein and encoding DNA
 PT used in gene therapy, is also useful for e.g. identifying ligands and
 PT raising antibodies and antisera -

XX PS Claim 1; Fig 1; 80pp; Japanese.

XX This represents a human brain-derived G-protein coupled receptor
 CC protein. The polypeptide can be expressed by standard recombinant
 CC methodology. The novel G-protein coupled receptor protein can be used
 CC for identifying ligands; raising antibodies and antisera; developing

CC receptor-binding assay system; and screening for drug candidates. The
 CC encoding polynucleotide can be used in; probing for diagnostic genes;
 CC constructing PCR primers; making transgenic animals; and in gene therapy.
 XX SQ Sequence 986 AA;

Query Match 78.0%; Score 32; DB 21; Length 986;
 Best Local Similarity 85.7%; Pred. NO. 3.6e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
 I I I I I
 Db 384 mlptylk 390

RESULT 9

AAW75457

ID AAW75457 standard; peptide; 8 AA.

XX AC AAW75457;

XX DT 27-APR-1999 (first entry)

XX DE Mammalian tub protein tyrosine phosphorylation site.

XX KW Human; wild type; tubby; identification; SH2 domain; mammal; obesity;
 KW body weight disorder; cachexia; anorexia.

XX OS Homo sapiens.

XX OS Mus sp.

XX PN US5861239-A.

XX PD 19-JAN-1999.

XX PF 02-SEP-1997; 97US-0922267.

XX PR 02-SEP-1997; 97US-0922267.

XX PR 12-APR-1996; 96US-0631200.

XX PR 28-MAR-1997; 97US-0829553.

XX PA (WILL-) MILLENNIUM PHARM INC.

XX PI Kapeller R, Kiehn PW, Moore KJ;

XX DR WPI; 1999-130383/11.

XX Identifying compounds which modulate tub protein activity - by
 PT detecting compounds which alter the interaction of tub protein with
 PT a SH2 containing peptide, used to develop agents for treating e.g.
 PT obesity, cachexia or anorexia

XX PS Disclosure; Column 13; 95pp; English.

XX This sequence represents a tyrosine phosphorylation site/SH2 docking
 CC domain found in the amino acid sequence of the mouse and human "tub"
 CC proteins (AAW75450 and AAW75451 respectively). The invention relates to
 CC a method for identifying compounds that modulate tub protein activity,
 CC especially its interaction with proteins containing an SH2 domain. The
 CC method can be used for identifying compounds which modulate tub protein
 CC activity for use in the treatment of mammalian body weight disorders
 CC including obesity, cachexia and anorexia.

XX SQ Sequence 8 AA;

Query Match 75.6%; Score 31; DB 20; Length 8;
 Best Local Similarity 100.0%; Pred. NO. 3.4e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTTY 6
 I I I I I

Db 2 mypty 6

RESULT 10

AAW97073
ID AAW97073 standard; peptide; 12 AA.

XX AC AAW97073;
XX DT 29-APR-1999 (first entry)

XX DE Peptidomimetic capable of inhibiting CD28 and/or CTLA-4 interactions.

XX KW Peptidomimetic; CD28; CTLA-4; CD80; CD86; B7-1; B7-2; diabetes;
XX KW immune system disease; autoimmune disease; psoriasis; multiple sclerosis;
XX KW lupus erythematosus; rheumatoid arthritis; transplant rejection; cancer.

XX OS Synthetic.

XX PN WO9856401-A1.

XX PD 17-DEC-1998.

XX PF 11-JUN-1998; 98WO-US12312.

XX PR 12-JUN-1997; 97US-0049470.

XX PA (ISTF) ARS APPLIED RES SYSTEMS HOLDING NV.

XX PI Blechner S, El Tayar N, Jameson B, Tepper M;

XX DR WPI; 1999-080854/07.

XX PT New peptidomimetic compounds - useful in the prophylaxis, diagnosis
XX PT and treatment of pathologies and disorders, which are improved by
XX PT inhibition of CD28 and/or CTLA-4 interaction with CD80 (B7-1) and
XX PT CD86 (B7-2)

XX PS Disclosure; Page 9; 62pp; English.

XX CC AAW97053-82 represent peptidomimetic compounds that inhibit CD28 and/or
XX CC CTLA-4 interactions with CD80 and CD86. The peptides are used in
XX CC pharmaceutical compositions for the prophylaxis, diagnosis and treatment
XX CC of pathologies and disorders, which are improved by inhibition of CD28
XX CC and/or CTLA-4 interaction with CD80 (B7-1) and CD86 (B7-2). Disorders
XX CC include immune system diseases (e.g. Autoimmune diseases such as
XX CC psoriasis, multiple sclerosis, lupus erythematosus, diabetes, rheumatoid
XX CC arthritis, and therapy against solid organ/cellular transplant
XX CC rejection) and cancer.

XX SQ Sequence 12 AA;

Query Match 75.6%; Score 31; DB 20; Length 12;

Best Local Similarity 71.4%; Pred No. 5.2;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 MYPTYLK 8

DB 2 myppylr 8

RESULT 11

AAAY12222
ID AAAY12222 standard; Protein; 58 AA.

XX AC AAAY12222;

XX DT 18-JUN-1999 (first entry)

XX DE Human 5' EST secreted protein SEQ ID NO: 535.

XX KW

KW

KW Human; secreted protein; EST; expressed sequence tag; diagnosis;
KW forensic; gene therapy; chromosome mapping; signal peptide;
KW upstream regulatory sequence; cytokine activity; cell proliferation;
KW differentiation; haematopoiesis regulation; tissue growth regulation;
KW thrombolytic; hormone regulation; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; anti-inflammatory; tumour inhibition.

XX OS Homo sapiens.

XX PN WO9906554-A2.

XX PD 11-FEB-1999.

XX PF 31-JUL-1998; 98WO-IB01238.

XX PR 01-AUG-1997; 97US-0905134.

XX PA (GEST) GENSET.

XX PI Duclert A, Dumas Milne Edwards J, Lacroix B;

XX DR WPI; 1999-153784/13.

XX PF N-PSDB; AAX41055.

XX PT New nucleic acids encoding human secreted proteins - obtained from
XX PT cDNA libraries prepared from kidney, fetal kidney, dystrophic
XX PT muscle, muscle and heart tissue

XX PS Claim 34; Page 598; 622pp; English.

XX CC AAX40826 to AAX41093 represent 5' expressed sequence tags (ESTs) for
XX CC human secreted proteins, and encode the proteins given in AAY01602 and
XX CC AAY11994 to AAY12260, respectively. The proteins given represent the
XX CC signal peptide and an N-terminal fragment of a secreted protein. The
XX CC nucleic acid sequences can be used for producing secreted human gene
XX CC products. They can also be used to develop products for diagnosis and
XX CC therapy. The proteins obtained may have cytokine activity, cell
XX CC proliferation/differentiation activity, haematopoiesis regulating
XX CC activity, tissue growth regulating activity, reproductive hormone
XX CC regulating activity, chemotactic/chemokinetic activity, haemostatic and
XX CC thrombolytic activity, receptor/ligand activity, anti-inflammatory
XX CC activity, tumour inhibition activity or other activities. The products
XX CC can be used in forensic, gene therapy and chromosome mapping procedures.
XX CC The sequences can also be used for obtaining corresponding promoter
XX CC sequences. The nucleic acids encoding the signal peptide can be used
XX CC for directing extracellular secretion of a polypeptide or the insertion
XX CC of a polypeptide into a membrane, or importing a polypeptide into
XX CC a cell.

XX SQ Sequence 58 AA;

Query Match 75.6%; Score 31; DB 20; Length 58;

Best Local Similarity 83.3%; Pred No. 28;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYL 7

DB 1 mypsyl 6

RESULT 12

AAAB74618
ID AAAB74618 standard; protein; 173 AA.

XX AC AAAB74618;

XX DT 22-MAY-2001 (first entry)

XX DE Human 72 kDa gelatinase catalytic domain protein sequence SEQ ID NO:8.

XX KW Human; 92 kDa gelatinase; 72 kDa gelatinase; matrilysin; gelatinase;
XX KW catalytic domain; 92 CD; 72 CD; matrilysin CD; elastin; vulnery;

KW excess connective tissue removal; dermatological; keloid; scleroderma;
 KW post-operative fibrosis; intervertebral disc injection; fibrotic disease;
 KW hypertrophic scar; wound debridement; post-surgical adhesion;
 XX idiopathic pulmonary fibrosis.

OS Homo sapiens.

XX US6194189-B1.

XX 27-FEB-2001.

XX 16-DEC-1994; 94US-0357820.

XX 16-DEC-1994; 94US-0357820.

XX (UNIW) UNIV WASHINGTON.

XX Senior RM;

XX WPI; 2001-243407/25.

XX New gelatinase truncated mutant useful for treating disorders requiring
 XX the removal of excess connective tissue, e.g. keloids, post-operative
 XX fibrosis, intervertebral disc injections, hypertrophic scars

XX Example; Column 13-16; l1pp; English.

XX The present invention describes a cDNA sequence which encodes the
 CC truncated mutant of the 92 kDa gelatinase having an amino acid sequence
 CC as given in AAB74617 consisting of residues 106-216 fused to residues
 CC 391-443 of the parent molecule. The truncated mutant has dermatological
 CC and vulnary activities. The truncated mutant is useful for treating
 CC disorders requiring the removal of excess connective tissue,
 CC e.g., keloids, post-operative fibrosis, intervertebral disc injections,
 CC hypertrophic scars, wound debridement, post-surgical adhesions and
 CC various fibrotic diseases (including scleroderma, idiopathic pulmonary
 CC fibrosis). The truncated mutant is catalytically active compared to
 CC the full protein. Unlike the full protein, the truncated mutant is
 CC essentially inactive against insoluble elastin, and does not require
 CC activation to be enzymatically active. The present sequence represents
 CC a 72 kDa gelatinase catalytic domain (72 CD) containing protein, which
 CC is given in the exemplification of the present invention.

XX Sequence 173 AA;

Query Match 75.6%; Score 31; DB 22; Length 173;
 Best Local Similarity 100.0%; Pred. No. 89;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6

Db 138 mypty 142

RESULT 13

AAB68682

ID AAB68682 standard; Protein; 173 AA.

XX AAB68682;

XX 02-MAY-2001 (first entry)

XX Matrilysin catalytic domain.

XX Gelatinase; excess connective tissue removal; matrilysin;
 KW matrix metalloproteinase.

XX Unidentified.

XX US6184021-B1.

XX 06-FEB-2001.

XX 19-MAY-1995; 95US-0444628.

XX 16-DEC-1994; 94US-0357820.

XX (UNIW) UNIV WASHINGTON.

XX Senior RM;

XX WPI; 2001-202001/20.

XX New truncated mutant of 92 kDa gelatinase which is catalytically
 PT active, but is inactive against insoluble elastin, useful for treating
 PT disorders requiring the removal of excess connective tissues such as
 PT keloids

XX Examples; Fig 2; l1pp; English.

XX The present invention relates to a truncated mutant (92 CD) of the 92
 CC kilo Dalton (kDa) gelatinase (see AAB68681). The truncated protein is
 CC useful for treating disorders requiring the removal of excess connective
 CC tissue, e.g. keloids, post-operative fibrosis, intervertebral disc
 CC injections, hypertrophic scars, wound debridement, post-surgical
 CC adhesions and various fibrotic diseases (scleroderma, idiopathic
 CC pulmonary fibrosis). Gelatinase is a matrix metalloproteinase and is also
 CC known as gelatinase B and MMP-9. The truncated protein is catalytically
 CC active comparable to the full protein but unlike the full protein is
 CC essentially inactive against insoluble elastin. The present sequence is
 CC the catalytic domain of matrilysin, which was used in a sequence homology
 CC alignment with the 92 CD protein of the present invention. Matrilysin is
 CC also a matrix metalloproteinase.

XX Sequence 173 AA;

Query Match 75.6%; Score 31; DB 22; Length 173;
 Best Local Similarity 100.0%; Pred. No. 89;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6

Db 138 mypty 142

RESULT 14

AAW76253

ID AAW76253 standard; Protein; 261 AA.

XX AAW76253;

XX 02-DEC-1998 (first entry)

XX Human matrilysin-like protein.

XX Matrilysin; inhibitor; tumour necrosis factor; screening; MMP; infection;
 KW matrix metalloproteinase; antagonist; diagnosis; underexpression; cancer;
 KW degeneration; extracellular matrix; arthritis; cardiovascular disease;
 KW cachexia; multiple sclerosis.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..22

FT Protein /label= signal

FT Protein 23..261

FT Protein /label= matrilysin_like_protein

XX WO9831818-A2.

XX 23-JUL-1998.

XX 20-JAN-1998; 98WO-US00783.

```

PR 01-AUG-1997; 97US-0054541.
PR 21-JAN-1997; 97US-0034205.
PR 13-JUN-1997; 97US-0049607.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Brower L, Gentz R, Ni J, Rosen CA, Ruben SM;
XX
XX WPI; 1998-414114/35.
XX
XX N-PSDB; AAV61633.
XX
XX Isolated nucleic acid encoding human metallo-protease(s) - used for
XX diagnosis, treatment and prevention of, e.g. cancer, inflammation,
XX neurological disease and infections
XX
XX Claim 20a; Fig 4; 8lpp; English.
XX
XX This sequence represents a novel human matrilysin-like protein which is
XX an inhibitor of the members of the matrix metalloproteinase, MMP,
XX family. This protein can be used in assays to screen for agonists and
XX antagonists and the nucleic acid is used as a probe for gene mapping, in
XX situ hybridisation and detection of corresponding genes in human tissue,
XX and as sources of probes and primers for diagnosis. The protein and its
XX antigenic fragments are used to raise antibodies (Ab) (which can be used
XX for diagnosis in usual immunoassays or for in vivo imaging) and to screen
XX for (ant)agonists. Agonists can be used to treat cancer or other
XX conditions associated with underexpression of TNF-alpha. Antagonists are
XX used to treat diseases associated with degeneration of the extracellular
XX matrix (cancer, arthritis, cardiovascular disease, cachexia and multiple
XX sclerosis).
XX
XX Sequence 261 AA;

Query Match 75.6%; Score 31; DB 19; Length 261;
Best Local Similarity 100.0%; Pred. NO. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
DB 226 mypty 230
|||||

RESULT 15
AAW76254
ID AAW76254 standard; Protein; 264 AA.
XX
XX AC AAW76254;
XX
XX 02-DEC-1998 (first entry)
XX
XX Human matrilysin protein.
XX
XX Matrilysin; inhibitor; tumour necrosis factor; screening; MMP; infection;
XX matrix metalloproteinase; antagonist; diagnosis; underexpression; cancer;
XX degeneration; extracellular matrix; arthritis; cardiovascular disease;
XX cachexia; multiple sclerosis.
XX
XX Homo sapiens.
XX
XX WO9831818-A2.
XX
XX 23-JUL-1998.
XX
XX 20-JAN-1998; 98WO-US00783.
XX
XX 01-AUG-1997; 97US-0054541.
XX 21-JAN-1997; 97US-0034205.
XX 13-JUN-1997; 97US-0049607.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Brower L, Gentz R, Ni J, Rosen CA, Ruben SM;

```

```

XX
XX WPI; 1998-414114/35.
XX
XX Isolated nucleic acid encoding human metallo-protease(s) - used for
XX diagnosis, treatment and prevention of, e.g. cancer, inflammation,
XX neurological disease and infections
XX
XX Disclosure; Fig 5; 8lpp; English.
XX
XX This sequence represents the human matrilysin protein which is an
XX inhibitor of the members of the matrix metalloproteinase, MMP, family.
XX This protein is used in the identification of a novel human
XX matrilysin-like protein which can be used in assays to screen for
XX agonists and antagonists and the nucleic acid is used as a probe for gene
XX mapping, in situ hybridisation and detection of corresponding genes in
XX human tissue, and as sources of probes and primers for diagnosis. The
XX protein and its antigenic fragments are used to raise antibodies (Ab)
XX (which can be used for diagnosis in usual immunoassays or for in vivo
XX imaging) and to screen for (ant)agonists. Agonists can be used to treat
XX cancer or other conditions associated with underexpression of TNF-alpha.
XX Antagonists are used to treat diseases associated with degeneration of
XX the extracellular matrix (cancer, arthritis, cardiovascular disease,
XX cachexia and multiple sclerosis).
XX
XX Sequence 264 AA;

Query Match 75.6%; Score 31; DB 19; Length 264;
Best Local Similarity 100.0%; Pred. NO. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
DB 232 mypty 236
|||||

Search completed: September 4, 2001, 15:50:25
Job time: 163 sec

```


THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2001, 15:51:12 : Search time 40.98 Seconds
(without alignments)
4.020 Million cell updates/sec

Title: US-09-630-345-4

Perfect score: 41

Sequence: 1 XMYPTYLK 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_AA:*
1: /cgn2.6/ptodata/2/1aa/5A_COMB.pep.*
2: /cgn2.6/ptodata/2/1aa/5B_COMB.pep.*
3: /cgn2.6/ptodata/2/1aa/6A_COMB.pep.*
4: /cgn2.6/ptodata/2/1aa/6B_COMB.pep.*
5: /cgn2.6/ptodata/2/1aa/PTUS_COMB.pep.*
6: /cgn2.6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	97.6	8	1 US-07-620-410-3	Sequence 3, Appli
2	31	75.6	8	2 US-08-922-267A-65	Sequence 65, Appl
3	31	75.6	173	4 US-08-444-628-8	Sequence 8, Appli
4	31	75.6	173	3 US-08-357-820-8	Sequence 8, Appli
5	31	75.6	261	3 US-09-009-156-5	Sequence 5, Appli
6	31	75.6	264	3 US-09-009-156-6	Sequence 6, Appli
7	31	75.6	267	4 US-08-448-489-18	Sequence 18, Appl
8	31	75.6	271	3 US-08-896-062-2	Sequence 2, Appli
9	31	75.6	459	1 US-08-630-592-2	Sequence 2, Appli
10	31	75.6	459	1 US-08-714-991-2	Sequence 2, Appli
11	31	75.6	459	3 US-09-032-365A-2	Sequence 2, Appli
12	31	75.6	460	3 US-08-630-592-7	Sequence 7, Appli
13	31	75.6	460	1 US-08-714-991-7	Sequence 7, Appli
14	31	75.6	460	3 US-09-032-365A-8	Sequence 8, Appli
15	31	75.6	462	4 US-08-068-392-3	Sequence 3, Appli
16	31	75.6	462	4 US-08-396-988-3	Sequence 3, Appli
17	31	75.6	505	1 US-08-631-200-2	Sequence 2, Appli
18	31	75.6	505	1 US-08-630-592-4	Sequence 4, Appli
19	31	75.6	505	1 US-08-714-991-4	Sequence 4, Appli
20	31	75.6	505	2 US-08-829-553-2	Sequence 2, Appli
21	31	75.6	505	2 US-08-922-267A-2	Sequence 2, Appli
22	31	75.6	505	2 US-08-936-707A-2	Sequence 2, Appli
23	31	75.6	505	2 US-08-936-706A-2	Sequence 2, Appli
24	31	75.6	505	3 US-09-248-203-2	Sequence 2, Appli
25	31	75.6	505	3 US-09-032-365A-4	Sequence 4, Appli
26	31	75.6	505	4 US-08-812-824-3	Sequence 3, Appli
27	31	75.6	505	4 US-09-406-071-2	Sequence 2, Appli

28 31 75.6 506 1 US-08-631-200-8 Sequence 8, Appli
29 31 75.6 506 2 US-08-829-553-8 Sequence 8, Appli
30 31 75.6 506 2 US-08-922-267A-8 Sequence 8, Appli
31 31 75.6 506 2 US-08-936-707A-8 Sequence 8, Appli
32 31 75.6 506 2 US-08-936-706A-8 Sequence 8, Appli
33 31 75.6 506 3 US-09-248-203-8 Sequence 8, Appli
34 31 75.6 506 3 US-09-032-365A-62 Sequence 62, Appli
35 31 75.6 506 4 US-08-812-824-4 Sequence 4, Appli
36 31 75.6 506 4 US-09-406-071-8 Sequence 8, Appli
37 31 75.6 512 3 US-09-032-365A-60 Sequence 60, Appli
38 31 75.6 518 3 US-09-032-365A-58 Sequence 27, Appli
39 31 75.6 561 1 US-08-714-991-27 Sequence 10, Appli
40 31 75.6 561 3 US-09-032-365A-10 Sequence 10, Appli
41 30 73.2 73 1 US-08-379-538-7 Sequence 7, Appli
42 30 73.2 452 2 US-08-731-079A-2 Sequence 2, Appli
43 30 73.2 547 4 US-09-178-252-11 Sequence 11, Appli
44 30 73.2 725 1 US-08-448-170-4 Sequence 4, Appli
45 30 73.2 725 4 US-08-961-803-7 Sequence 7, Appli

ALIGNMENTS

RESULT 1
US-07-620-410-3
: Sequence 3, Application US/07620410
: Patent No. 5217953
: GENERAL INFORMATION:
: APPLICANT: Gozes, Iilana
: APPLICANT: Brennenman, Douglas E.
: APPLICANT: Fridkin, Mati
: APPLICANT: Moody, Terry
: TITLE OF INVENTION: A NOVEL VASOACTIVE INTESTINAL PEPTIDE
: TITLE OF INVENTION: ANTAGONIST
: NUMBER OF SEQUENCES: 3
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
: STREET: 1615 L. Street, N.W.
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20036
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/620,410
: FILING DATE: 19901130
: CLASSIFICATION: 514
: ATTORNEY/AGENT INFORMATION:
: NAME: Scott, Watson T.
: REGISTRATION NUMBER: 26,581
: REFERENCE/DOCKET NUMBER: WTS/5683/82679
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202) 861-3000
: TELEFAX: (202) 822-0944
: TELEX: 248453 CUSH
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 8 amino acids
: TYPE: AMINO ACID
: TOPOLOGY: linear
US-07-620-410-3

Query Match 97.6%; Score 40; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
|||||||

Db 2 MYPTLK 8

RESULT 2
US-08-922-267A-65
; Sequence 65, Application US/08922267A
; Patent No. 5861239
; GENERAL INFORMATION:
; APPLICANT: Kleya, Patrick W.
; APPLICANT: Moore, Karen J.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; NUMBER OF SEQUENCES: 82
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/922,267A
; FILING DATE: 2-SEP-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/829,553
; FILING DATE: 28-MAR-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/631,200
; FILING DATE: 12-APR-1996
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-085
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-922-267A-65

Query Match 75.6%; Score 31; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 MYPTY 6
Db 2 MYPTY 6

RESULT 3
US-08-444-628-8
; Sequence 8, Application US/08444628
; Patent No. 6184021
; GENERAL INFORMATION:
; APPLICANT: Senior, Robert M.
; TITLE OF INVENTION: Catalytically-active Gelatinase Mutant
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Scott J. Meyer, Monsanto/Searle, A3SG
; STREET: 800 N. Lindbergh Blvd.
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63167
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/444,628
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/357,820
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyer, Scott J.
; REGISTRATION NUMBER: 25,275
; REFERENCE/DOCKET NUMBER: WU-2855
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)694-3117
; TELEFAX: (314)694-5435
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 173 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-444-628-8

Query Match 75.6%; Score 31; DB 4; Length 173;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 MYPTY 6
Db 138 MYPTY 142

RESULT 4
US-08-357-820-8
; Sequence 8, Application US/08357820
; Patent No. 6194189
; GENERAL INFORMATION:
; APPLICANT: Senior, Robert M.
; TITLE OF INVENTION: Catalytically-active Gelatinase Mutant
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scott J. Meyer, Monsanto/Searle, A3SG
; STREET: 800 N. Lindbergh Blvd.
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63167
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/357,820
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyer, Scott J.
; REGISTRATION NUMBER: 25,275
; REFERENCE/DOCKET NUMBER: WU-2855
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (314)694-3117
TELEFAX: (314)694-5435
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 173 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-357-820-8

Query Match 75.6%; Score 31; DB 4; Length 173;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
Db 138 MYPTY 142

RESULT 5
US-09-009-156-5
Sequence 5, Application US/09009156
Patent No. 6046031
GENERAL INFORMATION:
APPLICANT: Ni, Jian
APPLICANT: Ruben, Steven
APPLICANT: Brewer, Laurie
APPLICANT: Gentz, Reiner
TITLE OF INVENTION: No. 6046031el Metalloproteinas
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: US
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/009,156
FILING DATE:
CLASSIFICATION: 424

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/034,205
FILING DATE: 21-JAN-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/049,607
FILING DATE: 13-JUN-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/054,541
FILING DATE: 01-AUG-1997
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PF376
TELEPHONE: 301-309-8504
TELEFAX: 301-309-8439
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 261 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-009-156-5

Query Match 75.6%; Score 31; DB 3; Length 261;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
Db 226 MYPTY 230

RESULT 6
US-09-009-156-6
Sequence 6, Application US/09009156
Patent No. 6046031
GENERAL INFORMATION:
APPLICANT: Ni, Jian
APPLICANT: Ruben, Steven
APPLICANT: Brewer, Laurie
APPLICANT: Gentz, Reiner
TITLE OF INVENTION: No. 6046031el Metalloproteinas
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: US
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/009,156
FILING DATE:
CLASSIFICATION: 424

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/034,205
FILING DATE: 21-JAN-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/049,607
FILING DATE: 13-JUN-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/054,541
FILING DATE: 01-AUG-1997
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PF376
TELEPHONE: 301-309-8504
TELEFAX: 301-309-8439
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 264 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-009-156-6

Query Match 75.6%; Score 31; DB 3; Length 264;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
Db 232 MYPTY 236

RESULT 7

US-08-448-489-18

; Sequence 18, Application US/08448489

; Patent No. 1840222

; GENERAL INFORMATION:

; APPLICANT: SEIKI, Motoharu

; APPLICANT: SHINAGAWA, Akira

; TITLE OF INVENTION: NOVEL METALLOPROTEINASE AND ENCODING DNA THEREFOR

; FILE REFERENCE: 15-290P

; CURRENT APPLICATION NUMBER: US/08/448,489

; CURRENT FILING DATE: 1995-06-07

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: Patent in Ver. 2.0

; SEQ ID NO 18

; LENGTH: 267

; TYPE: PRT

; ORGANISM: Unknown

; FEATURE:

; OTHER INFORMATION: Description of Unknown Organism: Known Member of

; OTHER INFORMATION: Matrix Metalloproteinase Family

US-08-448-489-18

Query Match

Best Local Similarity 75.6%; Score 31; DB 4; Length 267;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6

Db 232 MYPTY 236

RESULT 8

US-08-896-062-2

; Sequence 2, Application US/08896062

; Patent No. 6010893

; GENERAL INFORMATION:

; APPLICANT: KIHARA, Yasunori

; TITLE OF INVENTION: PROCESS FOR PRODUCING HUMAN MATRILYSIN

; NUMBER OF INVENTION: BY MEANS OF RECOMBINANT DNA

; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Browdy and Neimark

; STREET: 419 Seventh Street N.W., Ste. 300

; CITY: Washington

; STATE: D.C.

; COUNTRY: United States of America

; ZIP: 20004

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/896,062

; FILING DATE: 17-JUL-1997

; CLASSIFICATION: 435

; PRIORITY INFORMATION:

; APPLICATION NUMBER: 08/530,984

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Browdy, Roger L.

; REGISTRATION NUMBER: 25,618

; REFERENCE/DOCKET NUMBER: KIHARA-2

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 628-5197

; TELEFAX: (202) 737-3528

; TELEX: 248633

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 271 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-896-062-2

Query Match

Best Local Similarity 75.6%; Score 31; DB 3; Length 271;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6

Db 236 MYPTY 240

RESULT 9

US-08-630-592-2

; Sequence 2, Application US/08630592

; Patent No. 5770432

; GENERAL INFORMATION:

; APPLICANT: Nishina, Patsy

; APPLICANT: No. 5770432enirauth, Konrad

; APPLICANT: No. 5770432enirauth, Michael

; TITLE OF INVENTION: Obesity Associated Genes

; NUMBER OF SEQUENCES: 25

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: FLEHR, HOBRACH, TEST, ALBRITTON & HERBERT

; STREET: 3400 Embarcadero Center, Suite 3400

; CITY: San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 941114187

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PCDOS/MSDOS

; SOFTWARE: Patent in Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/630,592

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Sherwood, Pamela J.

; REGISTRATION NUMBER: 36,677

; REFERENCE/DOCKET NUMBER: A59504/BIR/PJS

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 7811989

; TELEFAX: (415) 3983249

; TELEX: 910 277299

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 459 amino acids

; TYPE: amino acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

US-08-630-592-2

Query Match

Best Local Similarity 75.6%; Score 31; DB 1; Length 459;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6

Db 233 MYPTY 237

RESULT 10

US-08-714-991-2

; Sequence 2, Application US/08714991

; Patent No. 5776762

; GENERAL INFORMATION:

; APPLICANT: NORTH, Michael

APPLICANT: NISHINA, Patsy
APPLICANT: No. 5776762en-Trauth, Konrad
APPLICANT: NAGERT, Juergen
TITLE OF INVENTION: OBESITY ASSOCIATED GENES
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/714,991
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: SHERWOOD, Pamela J.
REGISTRATION NUMBER: 36,677
REFERENCE/DOCKET NUMBER: A-59504-1/PJS
TELEPHONE: 415-494-8700
TELEFAX: 415-494-8771
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 459 amino acids
TYPE: amino acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-714-991-2

Query Match 75.6%; Score 31; DB 1; Length 459;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
|||||
Db 233 MYPTY 237

RESULT 11
US-09-032-365A-2
Sequence 2, Application US/09032365A
Patent No. 6114502
GENERAL INFORMATION:
APPLICANT: No. 6114502th, Michael
APPLICANT: Nishina, Patsy
APPLICANT: NAGERT, Juergen
APPLICANT: No. 6114502en-Trauth, Konrad
TITLE OF INVENTION: GENE FAMILY ASSOCIATED WITH
TITLE OF INVENTION: NEUROSENSORY DEFECTS
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bozicevic & Reed, LLP
STREET: 285 Hamilton Avenue, Suite 200
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/032,365A
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Sherwood, Pamela J.
REGISTRATION NUMBER: 36,677
REFERENCE/DOCKET NUMBER: SEQ-2CIP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-327-3400
TELEFAX: 650 327-3231
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 459 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-032-365A-2

Query Match 75.6%; Score 31; DB 3; Length 459;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
|||||
Db 233 MYPTY 237

RESULT 12
US-08-630-592-7
Sequence 7, Application US/08630592
Patent No. 5770432
GENERAL INFORMATION:
APPLICANT: Nishina, Patsy
APPLICANT: No. 5770432enTrauth, Konrad
APPLICANT: NAGERT, Juergen
APPLICANT: No. 5770432th, Michael
TITLE OF INVENTION: Obesity Associated Genes
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
STREET: 3400 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 941114187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PCDOS/MSDOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/630,592
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sherwood, Pamela J.
REGISTRATION NUMBER: 36,677
REFERENCE/DOCKET NUMBER: A59504/BIR/PJS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 7811989
TELEFAX: (415) 3983249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 460 amino acids
TYPE: amino acid
STRANDEDNESS: double

TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-630-592-7

Query Match 75.6%; Score 31; DB 1; Length 460;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 MYPTY 6
Db 234 MYPTY 238

RESULT 13
US-08-714-991-7
Sequence 7, Application US/08714991
Patent No. 5776762
GENERAL INFORMATION:
APPLICANT: NORTH, Michael
APPLICANT: NISHINA, Patsy
APPLICANT: NAGGERT, Juergen
TITLE OF INVENTION: OBESITY ASSOCIATED GENES
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/714.991
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: SHERWOOD, Pamela J.
REGISTRATION NUMBER: 36,677
REFERENCE/DOCKET NUMBER: A-59504-1/PJS
TELEPHONE: 415-494-8700
TELEFAX: 415-494-8771
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 460 amino acids
TYPE: amino acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-714-991-7

Query Match 75.6%; Score 31; DB 1; Length 460;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 MYPTY 6
Db 234 MYPTY 238

RESULT 14
US-09-032-365A-8
Sequence 8, Application US/09032365A
Patent No. 6114502
GENERAL INFORMATION:

APPLICANT: No. 6114502th, Michael
APPLICANT: Nishina, Patsy
APPLICANT: Naggart, Juergen
APPLICANT: No. 6114502en-Trauth, Konrad
TITLE OF INVENTION: GENE FAMILY ASSOCIATED WITH
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bozicevic & Reed, LLP
STREET: 285 Hamilton Avenue, Suite 200
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/032.365A
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Sherwood, Pamela J.
REGISTRATION NUMBER: 36,677
REFERENCE/DOCKET NUMBER: SEQ-2C1P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-327-3400
TELEFAX: 650 327-3231
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 460 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-032-365A-8

Query Match 75.6%; Score 31; DB 3; Length 460;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 MYPTY 6
Db 234 MYPTY 238

RESULT 15
US-08-068-392-3
Sequence 3, Application US/08068392
Patent No. 6150152
GENERAL INFORMATION:
APPLICANT: Shapiro, Steven M.
TITLE OF INVENTION: Human Macrophage Metalloproteinase
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scott J. Meyer, Monsanto Co., A3SM
STREET: 800 N. Lindbergh Blvd.
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63160
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/068,392
;; FILING DATE: 19930528
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Meyer, Scott J.
;; REGISTRATION NUMBER: 25275
;; REFERENCE/DOCKET NUMBER: 07-24(12406)A
;; TELEPHONE: (314)694-3117
;; TELEFAX: (314)694-5435
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 462 amino acids
;; TYPE: AMINO ACID
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-08-068-392-3

Query Match 75.6%; Score 31; DB 4; Length 462;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTI 6
Db 229 MYPTI 233

Search completed: September 4, 2001, 15:51:13
Job time: 196 sec

THIS PAGE BLANK (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2001, 15:52:08 ; Search time 49.08 Seconds
(without alignments)
12.416 Million cell updates/sec

Title: US-09-630-345-4

Perfect score: 41

Sequence: 1 MYPTYLK 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_68.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37	90.2	384	2	collagenase (EC 3.1.1.1)
2	35	85.4	393	2	transporter ybfD [
3	34	82.9	436	2	4-aminobutyrate am
4	33	80.5	218	2	hypothetical prote
5	33	80.5	269	2	hypothetical prote
6	32	78.0	340	1	cell fusion protei
7	32	78.0	349	2	peroxidase (EC 1.1
8	32	78.0	593	2	organic cation tra
9	32	78.0	1263	2	DNA-directed RNA p
10	31	75.6	247	2	hypothetical prote
11	31	75.6	267	2	hypothetical prote
12	31	75.6	267	1	matrilysin (EC 3.4
13	31	75.6	267	2	matrilysin (EC 3.4
14	31	75.6	301	2	SEC14 protein - ye
15	31	75.6	378	2	hypothetical prote
16	31	75.6	385	2	Similar to seed ma
17	31	75.6	462	2	macrophage elastas
18	31	75.6	478	2	phospho-beta-gluc
19	31	75.6	483	2	matrix metalloprot
20	31	75.6	495	1	L1 protein - bovin
21	31	75.6	501	1	L1 protein - bovin
22	31	75.6	501	1	L1 protein - bovin
23	31	75.6	502	2	hypothetical prote
24	31	75.6	503	2	hypothetical prote
25	31	75.6	505	2	tub protein, brain
26	31	75.6	511	2	hypothetical prote
27	31	75.6	513	1	L1 protein - deer
28	31	75.6	516	2	hypothetical prote
29	31	75.6	565	2	hypothetical prote

30	31	75.6	562	2	S42826	probable ATPase -
31	31	75.6	2185	1	GNNYSV	genome polyprotein
32	31	75.6	2185	1	GNNYSV	genome polyprotein
33	30	73.2	141	2	T29507	hypothetical prote
34	30	73.2	152	2	G83476	hypothetical prote
35	30	73.2	155	2	T17986	hypothetical prote
36	30	73.2	192	2	T11930	NADH dehydrogenase
37	30	73.2	239	2	B64757	glycolate oxidase
38	30	73.2	239	2	D85522	probable dehydroge
39	30	73.2	248	2	S57910	probable response
40	30	73.2	262	2	C81384	shikimate 5-dehydr
41	30	73.2	287	2	D83414	probable transcrip
42	30	73.2	343	1	MMBEA5	cell fusion protei
43	30	73.2	368	1	HLBECM	membrane glycoprot
44	30	73.2	410	2	T20397	hypothetical prote
45	30	73.2	491	2	JE0396	phospho-beta-galac

ALIGNMENTS

RESULT 1
I51267
collagenase (EC 3.4.24.-) - bullfrog
C:Species: Rana catesbeiana (bullfrog)
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 22-Jun-1999
C:Accession: I51267
R:Oofusa, K.; Yomori, S.; Yoshizato, K.
Int. J. Dev. Biol. 38, 345-350, 1994
A:Title: Regionally and hormonally regulated expression of genes of collagen and coll
A:Reference number: I51267; MUID:95071832
A:Accession: I51267
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-384 <OOP>
A:Cross-references: GB:S75623; NID:g913070; PIDN:AAB32661.1; PID:g913071
C:Superfamily: interstitial collagenase; hemopexin repeat homology; matrix metallopro
C:Keywords: hydrolase; metalloproteinase; zinc; zymogen
F:53-231/Domain: matrix metalloproteinase homology <MMP>
F:236-381/Domain: hemopexin repeat homology <PN>
F:81-189,193,199/Binding site: zinc, catalytic (Cys, His, His, His) (inhibited) #stat
F:189,193,199/Binding site: zinc, catalytic (His) (active) #status predicted
F:190/Active site: Glu #status predicted

Query Match 90.2%; Score 37; DB 2; Length 384;
Best Local Similarity 85.7%; Pred. No. 5;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
|||||
DB 207 MYPTYL 213

RESULT 2
B86644
transporter ybfD [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 23-Mar-2001
C:Accession: B86644
R:Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarre, K.; Weissbach, J.; Eh
Genome Res. in press, 2001
A:Title: The complete genome sequence of the lactic acid bacterium.
A:Reference number: A86625
A:Accession: B86644
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-393 <STO>
A:Cross-references: GB:AE005176; NID:g12723004; PIDN:AAK04252.1; GSPDB:GN00146
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: ybfD

```

Query Match      85.4%; Score 35; DB 2; Length 393;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYL 7
DB 278 MYPTYL 283

RESULT 3
C69764
4-aminobutyrate aminotransferase homolog ycnG - Bacillus subtilis
C:Species: Bacillus subtilis
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
C:Accession: C69764
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bortol-
A.; Ehrlich, S.D.; Emerson, P.T.; Enlian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallen-
A.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F.;
A:Authors: Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
Y. M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle,
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron-
A.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MUID:98044033
A:Accession: C69764
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-436 <KUN>
A:Cross-references: GB:299106; GB:AL009126; NID:g2632653; PIDN:CAB12198.1; PID:g2632691
A:Experimental source: strain 168
C:Genetics:
A:Gene: ycnG
C:Superfamily: ornithine--oxo-acid aminotransferase

Query Match      82.9%; Score 34; DB 2; Length 436;
Best Local Similarity 71.4%; Pred. No. 23;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYL 8
DB 88 MYPTYIE 94

RESULT 4
E86750
hypothetical protein ykdB [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C:Species: Lactococcus lactis subsp. lactis
C>Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 23-Mar-2001
C:Accession: E86750
R:Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrlich
A:Title: The complete genome sequence of the lactic acid bacterium.
A:Reference number: A86625
A:Accession: E86750
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-218 <STO>
A:Cross-references: GB:AE005176; NID:gl2723952; PIDN:AAK05103.1; GSPDB:GN00146
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: ykdB

Query Match      80.5%; Score 33; DB 2; Length 218;
Best Local Similarity 78.0%; Score 32; DB 1; Length 340;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYL 8
DB 244 LYPTVIR 250

RESULT 5
D69011
hypothetical protein MTH1087 - Methanobacterium thermoautotrophicum (strain Delta H)
C:Species: Methanobacterium thermoautotrophicum
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C:Accession: D69011
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.
Q.; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwan,
A.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: fu
A:Reference number: A69000; MUID:98037514
A:Accession: D69011
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-269 <MTH>
A:Cross-references: GB:AE000879; GB:AE000666; NID:g2622175; PIDN:AA85576.1; PID:g262
A:Experimental source: strain Delta H
C:Genetics:
A:Gene: MTH1087

Query Match      80.5%; Score 33; DB 2; Length 269;
Best Local Similarity 83.3%; Pred. No. 23;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 YPTYLK 8
DB 175 YPTYIK 180

RESULT 6
MBE5
cell fusion protein precursor - human herpesvirus 3
C:Species: human herpesvirus 3, varicella-zoster virus
C>Date: 30-Sep-1988 #sequence_revision 30-Sep-1988 #text_change 16-Jul-1999
C:Accession: E27212
R:Davidson, A.J.; Scott, J.E.
J. Gen. Virol. 67, 1759-1816, 1986
A:Title: The complete DNA sequence of varicella-zoster virus.
A:Reference number: A27345; MUID:86306657
A:Accession: E27212
A:Molecule type: DNA
A:Residues: 1-340 <DAV>
A:Cross-references: EMBL:X04370; NID:g59989; PIDN:CAA27888.1; PID:g59994
C:Genetics:
A:Gene: 5
C:Superfamily: herpesvirus cell fusion protein
C:Keywords: membrane fusion; transmembrane protein
F:1-31/Domain: signal sequence #status predicted <SIG>
F:32-340/Product: cell fusion protein #status predicted <CFP>
F:115-137/Domain: transmembrane #status predicted <TM1>
F:220-236/Domain: transmembrane #status predicted <TM3>
F:251-269/Domain: transmembrane #status predicted <TM4>
F:307-322/Domain: transmembrane #status predicted <TM5>

Query Match      78.0%; Score 32; DB 1; Length 340;
Best Local Similarity 57.1%; Pred. No. 47;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYL 8
DB 244 LYPTVIR 250

```



```

RESULT 7
JU0458
peroxidase (EC 1.11.1.7) E - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 04-Mar-2000
C:Accession: JU0458
R:Intaprak, C.; Higashimura, N.; Yamamoto, K.; Okada, N.; Shimmyo, A.; Takano, M.
Gene 98, 237-241, 1991
A:Title: Nucleotide sequences of two genomic DNAs encoding peroxidase of Arabidopsis th
A:Reference number: JU0457; MUID:91200671
A:Accession: JU0458
A:Molecule type: DNA
A:Residues: 1-349 <INT>
A:Cross-references: GB:M58381; NID:gl66806; PIDN:AAA32842.1; PID:gl66807
C:Genetics:
A:Gene: prxEa
A:Introns: 76/3; 140/3; 197/1
A:Superfamily: peroxidase
C:Keywords: chromoprotein; heme; iron; metalloprotein; oxidoreductase
F:40-120/Disulfide bonds: #status predicted
F:67/Active site: Arg #status predicted
F:71,199/Binding site: heme iron (His) (axial ligands) #status predicted
F:73-78/Disulfide bonds: #status predicted
F:126-329/Disulfide bonds: #status predicted
F:206-238/Disulfide bonds: #status predicted

Query Match 78.0%; Score 32; DB 2; Length 349;
Best Local Similarity 83.3%; Pred. No. 48;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYL 7
:|||||
Db 226 LYPTYL 231

RESULT 8
JC4884
organic cation transporter protein 2 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 10-Sep-1996 #sequence_revision 18-Oct-1996 #text_change 05-Nov-1999
C:Accession: JC4884
R:Okuda, M.; Saito, H.; Urakami, Y.; Takano, M.; Inui, K.
Biochem. Biophys. Res. Commun. 224, 500-507, 1996
A:Title: cDNA cloning and functional expression of a novel rat kidney organic cation tra
A:Reference number: JC4884; MUID:96295517
A:Accession: JC4884
A:Molecule type: mRNA
A:Residues: 1-593 <OKU>
A:Cross-references: DBJ:D83044; NID:gl502282; PIDN:BAA11754.1; PID:dl012421; PID:gl5022
A:Experimental source: kidney
C:Comment: This protein is responsible for the transport of cationic drugs in kidney.

Query Match 78.0%; Score 32; DB 2; Length 593;
Best Local Similarity 57.1%; Pred. No. 82;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
:|||||
Db 457 LYPTYL 463

RESULT 9
F44466
DNA-directed RNA polymerase (EC 2.7.7.6) beta chain - Thermotoga maritima (strain MSB8)
C:Species: Thermotoga maritima
C:Date: 10-Jun-1993 #sequence_revision 26-Apr-1996 #text_change 21-Jul-2000
C:Accession: S41466; S41462; F44466; B72373; S19903
R:Palmer, P.; Schleper, C.; Arnold-Ammer, I.; Holz, I.; Meler, T.; Lottspeich, F.; Zillig,
Nucleic Acids Res. 21, 4904-4908, 1993
A:Title: The DNA-dependent RNA-polymerase of Thermotoga maritima; characterisation of th

```

```

A:Reference number: S41462; MUID:94232816
A:Accession: S41466
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1263 <PAL>
A:Cross-references: EMBL:X72695; NID:g425255; PIDN:CAA51246.1; PID:g425258
A:Experimental source: DSM 3109
A>Note: the nucleotide sequence was submitted to the EMBL Data Library; March 1993
A:Accession: S41462
A:Molecule type: protein
A:Residues: 1-5, 'X', 7-23 <PAW>
A:Experimental source: DSM 3109
R:Liao, D.; Dennis, P.P.
J. Biol. Chem. 267, 22787-22797, 1992
A:Title: The organization and expression of essential transcription translation compo
A:Reference number: A44466; MUID:93054590
A:Accession: F44466
A:Molecule type: DNA
A:Residues: 1-404 <LIA>
A:Cross-references: EMBL:X11839; NID:g407020; PIDN:CAA77863.1; PID:g48188
A>Note: sequence extracted from NCBI backbone (NCBIP:118059)
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.;
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson,
C.M.
Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome
A:Reference number: A72200; MUID:99287316
A:Accession: B72373
A:Molecule type: DNA
A:Residues: 1-1263 <ARN>
A:Cross-references: GB:AE001724; GB:AE000512; NID:g4980966; PIDN:AAD35543.1; PID:g498
A:Experimental source: strain MSB8
C:Genetics:
A:Gene: rpoB
A:Superfamily: DNA-directed RNA polymerase beta chain
C:Keywords: nucleotidyltransferase; transcription
F:1-1263/Product: DNA-directed RNA polymerase beta chain #status predicted <MAT>

Query Match 78.0%; Score 32; DB 2; Length 1263;
Best Local Similarity 83.3%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYL 7
:|||||
Db 214 LYPTYL 219

RESULT 10
B85515
hypothetical protein 20321 [imported] - Escherichia coli (strain O157:H7)
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 31-Mar-2001
C:Accession: B85515
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May
Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apoda
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: B85515
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-247 <STO>
A:Cross-references: GB:AE005174; NID:gl2513034; PIDN:AAG54582.1; GSPDB:GN00145; UNGP:
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: 20321
C:Superfamily: hypothetical protein f242a

Query Match 75.6%; Score 31; DB 2; Length 247;
Best Local Similarity 83.3%; Pred. No. 54;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY 3 YPTYLK 8
 Db 76 YPSYLK 81

RESULT 11
 T29698
 hypothetical protein F31A3.2 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
 C:Accession: T29698
 R:Murray, J.; Le, T.T.
 submitted to the EMBL Data Library, May 1996
 A:Description: The sequence of C. elegans cosmid F31A3.
 A:Reference number: Z20667
 A:Accession: T29698
 A:Status: Preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-262 <MUR>
 A:Cross-references: EMBL:U08742; PIDN:AAB36855.1; GSPDB:GN00028; CESP:F31A3.2
 A:Experimental source: strain Bristol N2; clone F31A3
 C:Genetics:
 A:Gene: CESP:F31A3.2
 A:Map position: X
 A:Introns: 23/3; 80/2; 155/2

Query Match 75.6%; Score 31; DB 2; Length 262;
 Best Local Similarity 71.4%; Pred. No. 58;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
 Db 188 MYPVYLE 194

RESULT 12
 KCHUM
 matrilysin (EC 3.4.24.23) precursor - human
 N:Alternate names: matrin; matrix metalloproteinase 7 (MMP7); probable metalloproteinase
 C:Species: Homo sapiens (man)
 C>Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 18-Jun-1999
 C:Accession: B28816; A60539; S24324
 R:Muller, D.; Quantin, B.; Gesnel, M.C.; Millon-Collard, R.; Abecassis, J.; Breathnach, B.
 Biochem. J. 253, 187-192, 1988
 A:Title: The collagenase gene family in humans consists of at least four members.
 A:Reference number: A90339; MUID:88339885
 A:Accession: B28816
 A:Molecule type: mRNA
 A:Residues: 1-267 <MUL>
 A:Cross-references: EMBL:X07819; NID:g935798; PIDN:CAA30678.1; PID:g35799
 R:Miyazaki, K.; Hattori, Y.; Umenishi, F.; Yasumitsu, H.; Umeda, M.
 Cancer Res. 50, 7758-7764, 1990
 A:Title: Purification and characterization of extracellular matrix-degrading metalloprotease from human placenta
 A:Reference number: A60539; MUID:91070531
 A:Accession: A60539
 A:Molecule type: Protein
 A:Residues: 18-35; X, 37-42 <MIY>
 R:Marci, H.P.; McNeill, L.; Thomas, G.; Davies, M.; Lovett, D.H.
 Biochem. J. 285, 899-905, 1992
 A:Title: Molecular characterization of a low-molecular-mass matrix metalloproteinase secreted from human placenta
 A:Reference number: S24324; MUID:92359961
 A:Accession: S24324
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-267 <MAR>
 A:Cross-references: EMBL:Z11887; NID:g35802; PIDN:CAA77942.1; PID:g35803
 C:Comments: This enzyme is similar in its activity to stromelysin and degrades various extracellular matrix components, including type I, IV, IX, X, and XI.
 C:Genetics: Matrilysin hydrolyzes peptide bonds in plasminogen to yield a fragment with a

A:Gene: GDB:MMP7; MPSLI
 A:Cross-references: GDB:125751; OMIM:178990
 A:Map position: 11q21-11q22
 C:Superfamily: matrilysin; matrix metalloproteinase homology
 C:Keywords: calcium; extracellular matrix; fibroblast; hydrolase; metalloproteinase;
 F:1-17/Domain: signal sequence #status predicted <SIG>
 F:18-267/Product: promatrilysin #status predicted <PRO>
 F:18-94/Domain: activation peptide #status predicted <ACT>
 F:55-259/Domain: matrix metalloproteinase homology <MMP>
 F:85-92/Region: autoinhibitory
 F:95-267/Product: matrilysin #status predicted <MAT>
 F:87,214,218,224/Binding site: zinc, catalytic (Cys, His, His) (inhibited) #stat
 F:214,218,224/Binding site: zinc, catalytic (His) (active) #status predicted
 F:215/Active site: Glu #status predicted

Query Match 75.6%; Score 31; DB 1; Length 267;
 Best Local Similarity 100.0%; Pred. No. 59;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
 Db 232 MYPTY 236

RESULT 13
 A57490
 matrilysin (EC 3.4.24.23) precursor - rat
 N:Alternate names: matrix metalloproteinase 7 (MMP7)
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 08-Dec-1995 #sequence_revision 08-Dec-1995 #text_change 22-Jun-1999
 C:Accession: A57490
 R:Abramson, S.R.; Conner, G.E.; Nagase, H.; Neuhaus, I.; Woessner Jr., J.F.
 J. Biol. Chem. 270, 16016-16022, 1995
 A:Title: Characterization of rat uterine matrilysin and its cDNA. Relationship to human
 A:Reference number: A57490; MUID:95332299
 A:Accession: A57490
 A:Status: Preliminary
 A:Molecule type: mRNA
 A:Residues: 1-267 <ABR>
 A:Cross-references: GB:L24374; NID:9402492; PIDN:AA99432.1; PID:g402493
 C:Superfamily: matrilysin; matrix metalloproteinase homology
 C:Keywords: hydrolase; metalloproteinase; zinc; zymogen
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-267/Product: matrilysin #status predicted <MAT>
 F:58-262/Domain: matrix metalloproteinase homology <MMP>
 F:90,217,221,227/Binding site: zinc, catalytic (Cys, His, His) (inhibited) #stat
 F:217,221,227/Binding site: zinc, catalytic (His) (active) #status predicted
 F:218/Active site: Glu #status predicted

Query Match 75.6%; Score 31; DB 2; Length 267;
 Best Local Similarity 100.0%; Pred. No. 59;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
 Db 235 MYPTY 239

RESULT 14
 S57923
 SEC14 protein - yeast (Candida albicans)
 C:Species: Candida albicans
 C>Date: 13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 26-Feb-1998
 C:Accession: S57923; S72193
 R:Montoliva, L.; Sanchez, M.; Pla, J.; Gil, C.; Nombela, C.
 submitted to the EMBL Data Library, September 1994
 A:Description: Characterisation of the Candida albicans SEC14 homolog gene.
 A:Reference number: S57923
 A:Accession: S57923
 A:Status: Preliminary
 A:Molecule type: DNA

A;Residues: 1-301 <MON>
 A;Cross-references: EMBL:X81937
 R;Monteoliva, L.; Sanchez, M.; Pla, J.; Gil, C.; Nombela, C.
 Yeast 12, 1097-1105, 1996
 A;Title: Cloning of Candida albicans SEC14 gene homologue coding for a putative essential
 A;Reference number: S72193; MUID:97051600
 A;Accession: S72193

A;Molecule type: DNA
 A;Residues: 1-301 <NOW>
 A;Cross-references: EMBL:X81937
 A;Note: the authors translated the codon CTG for residue 180 as Ser
 C;Genetics:
 A;Gene: SEC14
 C;Superfamily: cellular retinaldehyde-binding protein; cellular retinaldehyde-binding pr
 F;59-266/Domain: cellular retinaldehyde-binding protein homology <CRB>

Query Match 75.6%; Score 31; DB 2; Length 301;
 Best Local Similarity 100.0%; Pred. No. 66;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
 |||||
 Db 106 MYPTY 110

RESULT 15
 F86307
 hypothetical protein AAD50017.1 [imported] - Arabidopsis thaliana
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C;Accession: F86307
 R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A;Reference number: A86141; MUID:21016719
 A;Accession: F86307
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-378 <STO>
 A;Cross-references: GB:AE005172; NID:g5734752; PIDN:AAD50017.1; GSPDB:GN00141
 C;Genetics:
 A;Map position: 1

Query Match 75.6%; Score 31; DB 2; Length 378;
 Best Local Similarity 100.0%; Pred. No. 84;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
 |||||
 Db 118 MYPTY 122

Search completed: September 4, 2001, 15:52:09
 Job time: 232 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2001, 15:54:09 ; Search time 27.19 Seconds
(without alignments)
10.079 Million cell updates/sec

Title: US-09-630-345-4
Perfect score: 41
Sequence: 1 XMYPTYLK 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37	90.2	384	MM01_RANCA	Q11133 rana catesb
2	34	82.9	436	GABT_BACSU	P94427 bacillus su
3	32	78.0	340	CELF_VZVD	P09261 varicella-z
4	32	78.0	349	PERE_ARATH	P24102 arabidopsis
5	32	78.0	594	NU5M_HIPAM	Q92291 hippopotamu
6	32	78.0	1263	RPOB_THEMA	P29398 thermotoga
7	31	75.6	262	MM07_FELCA	P55032 felis silve
8	31	75.6	264	MM07_MOUSE	Q10738 mus musculu
9	31	75.6	267	MM07_HUMAN	P09237 homo sapien
10	31	75.6	267	MM07_RAT	P50280 rattus norv
11	31	75.6	301	SC14_CANAL	P46250 candida alb
12	31	75.6	462	MM12_MOUSE	P34960 mus musculu
13	31	75.6	465	MM12_MOUSE	Q63341 rattus norv
14	31	75.6	467	MM18_XENLA	Q13065 xenopus lae
15	31	75.6	482	MM20_MOUSE	P57748 mus musculu
16	31	75.6	483	MM20_MOUSE	Q60882 homo sapien
17	31	75.6	483	MM20_PTG	P79287 sus scrofa
18	31	75.6	495	VLI1_BPV1	P03103 bovine papi
19	31	75.6	497	VLI1_BPV2	P06458 bovine papi
20	31	75.6	501	VLI1_PAPVE	P11326 european el
21	31	75.6	503	YPT5_CAEEL	Q23469 caenorhabdi
22	31	75.6	505	TUB_MOUSE	P50586 mus musculu
23	31	75.6	505	TUB_RAT	O88808 rattus norv
24	31	75.6	506	TUB_HUMAN	P50607 homo sapien
25	31	75.6	511	YEL1_CAEEL	P90756 caenorhabdi
26	31	75.6	513	VLI1_PAPVD	P03104 deer papill
27	31	75.6	662	YME1_SCHMA	P46508 schistosoma
28	31	75.6	2185	POLG_SVDVH	P16604 s genome po
29	31	75.6	2185	POLG_SVDVU	P13900 s genome po
30	30	73.2	91	Y13F_BPT4	P39499 bacterioph
31	30	73.2	192	NU6M_PROWI	Q37622 prototheca
32	30	73.2	239	YKGE_ECOLI	P77252 eschericha
33	30	73.2	308	Y04O_BPT4	P39254 bacterioph

34	30	73.2	343	1	CELF_HSVEB	P28933 equine herp
35	30	73.2	368	1	VGH3_HCMVA	P08560 human cytom
36	30	73.2	425	1	BIOA_SERMA	P36568 serratia ma
37	30	73.2	580	1	P69_MYCHR	P15362 mycoplasma
38	30	73.2	634	1	VC09_VACCC	P21042 vaccinia vi
39	30	73.2	634	1	VC09_VACCV	P17372 vaccinia vi
40	30	73.2	1056	1	DPOL_ADE02	P03261 human adeno
41	30	73.2	1056	1	DPOL_ADE05	P04495 human adeno
42	30	73.2	1138	1	C7AA_BACTU	Q03749 bacillus th
43	30	73.2	1138	1	C7AA_BACUA	Q45707 bacillus th
44	30	73.2	1138	1	C7AB_BACUK	Q45708 bacillus th
45	30	73.2	1155	1	CIAB_BACTK	P06578 bacillus th

ALIGNMENTS

RESULT	1	MM01_RANCA	STANDARD;	PRT;	384 AA.
ID	MM01_RANCA				
AC	Q11133;				
DT	01-OCT-1996 (Rel. 34, Created)				
DT	01-OCT-1996 (Rel. 34, Last sequence update)				
DT	01-OCT-2000 (Rel. 40, Last annotation update)				
DE	INTERSTITIAL COLLAGENASE PRECURSOR (EC 3.4.24.7) (MATRIX				
DE	METALLOPROTEINASE-1) (MMP-1) (TCL).				
OS	Rana catesbeiana (Bull frog).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Amphibia; Batrachia; Anura; Neobatrachia; Ranidae; Rana.				
OX	NCBI_TaxID=8400;				
[1]					
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Skin;				
RX	MEDLINE=95071832; PubMed=7981043;				
RA	Oofusa K., Yomori S., Yoshizato K.;				
RT	"Regionally and hormonally regulated expression of genes of collagen				
RT	and collagenase in the anuran larval skin.";				
RL	Int. J. Dev. Biol. 38:345-350(1994).				
CC	-!- FUNCTION: CLEAVES COLLAGENS OF TYPES I, II, AND III AT ONE SITE IN				
CC	THE HELICAL DOMAIN. ALSO CLEAVES COLLAGENS OF TYPES VII AND X.				
CC	-!- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY.				
CC	-!- ENZYME REGULATION: CAN BE ACTIVATED WITHOUT REMOVAL OF THE				
CC	ACTIVATION PEPTIDE (BY SIMILARITY).				
CC	-!- SIMILARITY: CONTAINS 1 HEMOPEXIN-LIKE DOMAIN.				
CC	-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC				
CC	METALLOPROTEASE) ALSO KNOWN AS MATRININ SUBFAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	-----				
CC	EMBL; S75623; AAB32661.1; -				
DR	HSSP; P21692; IPBL.				
DR	MEROPS; M10.001; -				
DR	InterPro; IPR000130; -				
DR	InterPro; IPR000585; -				
DR	InterPro; IPR001818; -				
DR	Pfam; PF00413; Peptidase_M10; 1.				
DR	Pfam; PF00045; hemopexin; 2.				
DR	PROSITE; PS00024; HEMOPEXIN; FALSE_NEG.				
DR	PROSITE; PS00142; ZINC_PROTEASE; 1.				
DR	PROSITE; PS00546; CYSTEINE_SWITCH; FALSE_NEG.				
KW	Hydrolase; Metalloprotease; zinc; zymogen; Calcium;				
KW	Collagen degradation; Extracellular matrix; Signal.				
FT	SIGNAL 1 25 POTENTIAL.				
FT	PROPEP 26 88 ACTIVATION PEPTIDE (POTENTIAL).				
FT	CHAIN 89 384 INTERSTITIAL COLLAGENASE.				
FT	DOMAIN 239 384 HEMOPEXIN-LIKE.				
FT	SITE 81 CYSTEINE SWITCH (POTENTIAL).				

```

FT METAL 189 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 190 BY SIMILARITY.
FT METAL 193 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 199 ZINC (CATALYTIC) (BY SIMILARITY).
FT DISULFID 242 381 PROBABLE.
SQ SEQUENCE 384 AA; 43582 MW; A5B5E2FB332239DF CRC64;

Query Match 90.2%; Score 37; DB 1; Length 384;
Best Local Similarity 85.7%; Pred. No. 2.9;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
DB 207 MYPTYLK 213

RESULT 2
GABT_BACSU STANDARD; PRT; 436 AA.
GABT_BACSU
P94427;
DT 15-DEC-1998 (Rel. 37, Created)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PROBABLE 4-AMINO BUTYRATE AMINOTRANSFERASE (EC 2.6.1.19) (GAMMA-AMINO-
DE N-BUTYRATE TRANSAMINASE) (GABA TRANSAMINASE) (GLUTAMATE-SUCCINIC
DE SEMIALDEHYDE TRANSAMINASE) (GABA AMINOTRANSFERASE) (GABA-AT).
GN GABT.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168; PubMed=8969502;
RX MEDLINE=97124189; Kumano M., Kurita K.;
RA Imane K.;
RT The 25 degrees-36 degrees region of the Bacillus subtilis
RT chromosome: determination of the sequence of a 146 kb segment and
RT identification of 113 genes.
RL Microbiology 142:3047-3056(1996).
CC -!- CATALYTIC ACTIVITY: 4-AMINO BUTANOATE + 2-OXOGLUTARATE = SUCCINATE
CC -!- SEMIALDEHYDE + L-GLUTAMATE.
CC -!- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -!- PATHWAY: 4-AMINO BUTYRATE (GABA) DEGRADATION PATHWAY.
CC -!- SIMILARITY: BELONGS TO CLASS-III OF PYRIDOXAL-PHOSPHATE-DEPENDENT
CC AMINOTRANSFERASES.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; D50453; BAA09021.1; -
DR EMBL; Z99106; CAB2198.1; -
DR Subtilisin; Bgl2043; gabt.
DR InterPro; IPR000994; -
DR Pfam; PF00202; aminotran_3; 1.
DR PROSITE; PS00600; AA_TRANSFER_CLASS_3; 1.
RW Transferrase; Aminotransferase; Pyridoxal phosphate.
FT BINDING 281 281 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ SEQUENCE 436 AA; 47249 MW; D0961F6D4189A8F3 CRC64;

Query Match 82.9%; Score 34; DB 1; Length 436;
Best Local Similarity 71.4%; Pred. No. 13;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
DB 207 MYPTYLK 213

```

```

DB 88 MYPTYLK 94

RESULT 3
CELFL_VZVD STANDARD; PRT; 340 AA.
ID CELFL_VZVD
AC P09261;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE CELL FUSION PROTEIN PRECURSOR.
GN 5.
OS Varicella-zoster virus (strain Dumas) (VZV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10338;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86306657; PubMed=3018124;
RA Davison A.J., Scott J.E.;
RT "The complete DNA sequence of varicella-zoster virus.";
RL J. Gen. Virol. 67:1759-1816(1986).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; X04370; CAA27888.1; -
DR PIR; E27212; MMBS.
DR InterPro; IPR002567; -
DR Pfam; PF01621; Fusion_gly_K; 1.
KW Fusion protein; Transmembrane; Signal.
FT SIGNAL 1 ?
FT CHAIN ? 340 CELL FUSION PROTEIN.
FT CARBOHYD 55 55 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 65 65 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 340 AA; 38576 MW; 0387FE0EC39C946 CRC64;

Query Match 78.0%; Score 32; DB 1; Length 340;
Best Local Similarity 57.1%; Pred. No. 25;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
DB 244 LYPTYLK 250

RESULT 4
PERE_ARATH STANDARD; PRT; 349 AA.
ID PERE_ARATH
AC P24102;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE BASIC PEROXIDASE E PRECURSOR (EC 1.11.1.7).
GN PRXA.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91200671; PubMed=2016063;
RA Intapurk C., Higashimura N., Yamamoto K., Okada N., Shimmyo A.,
RA Takano M.;
RT "Nucleotide sequences of two genomic DNAs encoding peroxidase of
Arabidopsis thaliana.";

```

RL Gene 98:237-241(1991).
 CC -1- FUNCTION: REMOVAL OF H(2)O(2), OXIDATION OF TOXIC REDUCTANTS,
 CC BIOSYNTHESIS AND DEGRADATION OF LIGNIN, DEFENSE RESPONSE TOWARD
 CC WOUNDING AND METABOLISM OF AUXIN. THESE FUNCTIONS MIGHT BE
 CC DEPENDENT ON EACH ISOZYME/ISOFORM IN EACH PLANT TISSUE.
 CC -1- CATALYTIC ACTIVITY: DONOR + H(2)O(2) - OXIDIZED DONOR + 2 H(2)O.
 CC COFACTOR: HEME.
 CC -1- TISSUE SPECIFICITY: ROOTS.
 CC -1- SIMILARITY: BELONGS TO THE PEROXIDASE FAMILY. PLANT PEROXIDASE
 CC SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M59381; AAA32842.1; -;
 DR PIR; J00458; J00458.
 DR HSSP; P00433; IATJ.
 DR InterPro; IPR000823; -;
 DR InterPro; IPR002016; -;
 DR Pfam; PF00141; peroxidase; 1.
 DR PRINTS; PR00458; PEROXIDASE.
 DR PRINTS; PR00461; PLPEROXIDASE.
 DR PROSITE; PS00435; PEROXIDASE.1; 1.
 DR PROSITE; PS00436; PEROXIDASE.2; 1.
 DR Oxidoreductase; Glycoprotein; Peroxidase; Heme; Multigene family;
 KW Signal.
 FT SIGNAL 1 19
 FT CHAIN 20 349 BASIC PEROXIDASE E.
 FT ACT_SITE 67 67 BY SIMILARITY.
 FT ACT_SITE 71 71 DISTAL HISTIDINE (BY SIMILARITY).
 FT ACT_SITE 199 199 PROXIMAL HISTIDINE (HEME AXIAL LIGAND).
 FT DISULFID 40 120 BY SIMILARITY.
 FT DISULFID 73 78 BY SIMILARITY.
 FT DISULFID 126 329 BY SIMILARITY.
 FT DISULFID 206 238 BY SIMILARITY.
 FT CARBOHYD 86 86 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 173 173 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 187 187 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 217 217 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 243 243 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 349 AA; 38172 MW; 6E02C5758C36AB8F CRC64;
 Query Match 78.0%; Score 32; DB 1; Length 349;
 Best Local Similarity 83.3%; Pred. No. 26;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 2 MYPTYL 7
 Db 226 LYPTYL 231
 RESULT 5
 NU5M_HIPAM STANDARD; PRT; 594 AA.
 AC Q92ZY1;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DE NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 5 (EC 1.6.5.3).
 GN MTND5 OR ND5 OR NADH5.
 OS Hippopotamus amphibius (Hippopotamus).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Hippopotamidae; Hippopotamus.
 OX NCBI_TaxID=9833;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Ursing B.M., Arnason U.;
 RT "Analyses of mitochondrial genomes strongly support a hippopotamus-
 RT whale clade."
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE - NAD(+) + UBIQUINOL.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AJ010957; CAA09438.1; -;
 DR InterPro; IPR001516; -;
 DR InterPro; IPR001750; -;
 DR Pfam; PF00361; oxidored_q1; 1.
 DR Pfam; PF00662; oxidored_q1_N; 1.
 DR Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
 KW SEQUENCE 594 AA; 66599 MW; 4130499096B5A5CE CRC64;
 SQ
 Query Match 78.0%; Score 32; DB 1; Length 594;
 Best Local Similarity 83.3%; Pred. No. 44;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 3 YPTYLK 8
 Db 32 YPTVVK 37
 RESULT 6
 RPOB_THEMA STANDARD; PRT; 1263 AA.
 ID RPOB_THEMA
 AC P29338;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE BETA-DIRECTED RNA POLYMERASE BETA CHAIN (EC 2.7.7.6) (TRANSCRIPTASE
 DE BETA CHAIN) (RNA POLYMERASE BETA SUBUNIT).
 GN RPOB OR TM0458.
 OS Thermotoga maritima.
 OC Bacteria; Thermotogales; Thermotoga.
 OX NCBI_TaxID=2336;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MSB8 / DSM 3109;
 RX MEDLINE=94232816; PubMed=8177738;
 RA Palm P., Schleper C., Arnold-Ammer I., Holz I., Meier T.,
 RA Lottspeich F., Zillig W.;
 RT "The DNA-dependent RNA-polymerase of Thermotoga maritima;
 RT characterisation of the enzyme and the DNA-sequence of the genes for
 RT the large subunits."
 RL Nucleic Acids Res. 21:4904-4908(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MSB8 / DSM 3109;
 RX MEDLINE=99287316; PubMed=10360571;
 RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
 RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
 RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
 RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
 RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
 RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
 RT "Evidence for lateral gene transfer between Archaea and Bacteria from
 RT genome sequence of Thermotoga maritima."
 RL Nature 399:323-329(1999).
 RN [3]
 RP SEQUENCE OF 1-404 FROM N.A.
 RC STRAIN=MSB8 / DSM 3109;
 RX MEDLINE=93054590; PubMed=1429627;
 RN Liao D., Dennis P.P.;

```

RT RT The organization and expression of essential transcription
RT RT Translational component genes in the extremely thermophilic eubacterium
RT RT Thermotoga maritima.
RL J. Biol. Chem. 267:22787-22797(1992).
CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
CC SUBSTRATES.
CC -1- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE + N PYROPHOSPHATE +
CC RNA(N).
CC -1- SURUNIT: THE ENZYME CONSISTS OF THE SIGMA CHAIN AND THE CORE
CC ENZYME WHICH IS COMPOSED OF 2 ALPHA CHAINS, 1 BETA CHAIN, AND 1
CC BETA' CHAIN.
CC -1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA CHAIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X72695; CAA51246.1; -.
CC EMBL; AE001724; AAD35543.1; -.
CC EMBL; Z11839; CAA77863.1; -.
CC PIR; S19903; S19903.
CC PIR; S41466; S41466.
CC TIGR; TM0458; -.
CC InterPro; IPR001572; -.
CC Pfam; PF00562; RNA_POL_B; 1.
CC PROSITE; PS01166; RNA_POL_BETA; 1.
CC Transference; Transcription; DNA-directed RNA polymerase.
CC SEQUENCE 1263 AA; 143137 MW; 04B79368567C8237 CRC64;
SQ
Query Match 78.0%; Score 32; DB 1; Length 1263;
Best Local Similarity 83.3%; Pred. No. 94;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 2 MYPTVL 7
DB 214 LPTVL 219
RESULT 7
MM07_FELCA
ID MM07_FELCA STANDARD; PRT; 262 AA.
P55032;
01-OCT-1996 (Rel. 34, Created)
01-OCT-1996 (Rel. 34, Last sequence update)
01-OCT-2000 (Rel. 40, Last annotation update)
DE MATRILYSIN PRECURSOR (EC 3.4.24.23) (PUMP-1 PROTEASE) (UTERINE
DE METALLOPROTEINASE) (MATRIX METALLOPROTEINASE-7) (MMP-7) (MATRIN)
DE (FRAGMENT).
GN MMP7.
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9685;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Endometrium.
RA Scalzo C.M., Verhage H.G., Jaffe R.C.;
RT "Expression and estrogen control of PUMP-1 mRNA in the cat uterus.";
RL Endocrinol. Jpn. 2:229-235(1994).
CC -1- FUNCTION: DEGRADATES CASEIN, GELATINS OF TYPES I, III, IV, AND V,
CC AND FIBRONECTIN. ACTIVATES PROCOLLAGENASE.
CC -1- CATALYTIC ACTIVITY: CLEAVAGE OF 14-ALA-1-LEU-15 AND 16-TYR-1-
CC LEU-17 IN B CHAIN OF INSULIN. NO ACTION ON COLLAGEN TYPES I, II,
CC IV, V. CLEAVES GELATIN CHAIN ALPHA-2(I) > ALPHA-1(I).
CC -1- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC
CC METALLOPROTEASE) ALSO KNOWN AS MATRININ SUBFAMILY.
CC

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U04444; AAA18222.1; -.
CC HSP; P09237; IMMR.
CC MEROPS; M10.008; -.
CC InterPro; IPR000130; -.
CC InterPro; IPR001818; -.
CC Pfam; PF00413; Peptidase_M10; 1.
CC PROSITE; PS00142; ZINC_PROTEASE; 1.
CC PROSITE; PS00546; CYSTEINE_SWITCH; FALSE_NEG.
CC Hydrolase; Metalloprotease; Zinc; Zymogen; Calcium;
CC Collagen degradation; Extracellular matrix; Signal.
CC NON_TER 1
CC SIGNAL <1 12 BY SIMILARITY.
CC FT PROPEP 13 89 ACTIVATION PEPTIDE (BY SIMILARITY).
CC FT CHAIN 90 262 MATRILYSIN.
CC FT SITE 82 82 CYSTEINE SWITCH (POTENTIAL).
CC FT METAL 209 209 ZINC (CATALYTIC) (BY SIMILARITY).
CC FT ACT_SITE 210 210 BY SIMILARITY.
CC FT METAL 213 213 ZINC (CATALYTIC) (BY SIMILARITY).
CC FT METAL 219 219 ZINC (CATALYTIC) (BY SIMILARITY).
CC SEQUENCE 262 AA; 29263 MW; B4A1FA23320DC732 CRC64;
SQ
Query Match 75.6%; Score 31; DB 1; Length 262;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 MYPTY 6
DB 227 MPTY 231
RESULT 8
MM07_MOUSE
ID MM07_MOUSE STANDARD; PRT; 264 AA.
Q10738;
01-OCT-1996 (Rel. 34, Created)
01-OCT-1996 (Rel. 34, Last sequence update)
01-OCT-2000 (Rel. 40, Last annotation update)
DE MATRILYSIN PRECURSOR (EC 3.4.24.23) (PUMP-1 PROTEASE) (UTERINE
DE METALLOPROTEINASE) (MATRIX METALLOPROTEINASE-7) (MMP-7) (MATRIN).
GN MMP7.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR; TISSUE=uterus;
RX MEDLINE=96086498; Pubmed=7579699;
RA Wilson C.L., Heppner K.J., Rudolph L.A., Matrisian L.M.;
RT "The metalloproteinase matrilysin is preferentially expressed by
RT epithelial cells in a tissue-restricted pattern in the mouse.";
RL Mol. Biol. Cell 6:851-869(1995).
CC -1- FUNCTION: DEGRADATES CASEIN, GELATINS OF TYPES I, III, IV, AND V,
CC AND FIBRONECTIN. ACTIVATES PROCOLLAGENASE (BY SIMILARITY).
CC -1- FUNCTION: MAY PLAY A ROLE IN TISSUE REORGANIZATION.
CC -1- CATALYTIC ACTIVITY: CLEAVAGE OF 14-ALA-1-LEU-15 AND 16-TYR-1-
CC LEU-17 IN B CHAIN OF INSULIN. NO ACTION ON COLLAGEN TYPES I, II,
CC IV, V. CLEAVES GELATIN CHAIN ALPHA-2(I) > ALPHA-1(I).
CC -1- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC
CC METALLOPROTEASE) ALSO KNOWN AS MATRININ SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```


KW Collagen degradation; Extracellular matrix; Signal; Polymorphism.
FT SIGNAL 1 17
FT PROPEP 18 94
FT CHAIN 95 267
FT SITE 87 87
FT METAL 214 214
FT ACT_SITE 215 215
FT METAL 218 218
FT METAL 224 224
FT VARIANT 77 77
SQ SEQUENCE 267 AA; 29677 MW; F6BD1FC0ADA23603 CRC64;
/FTID=VAR_006729.

Query Match 75.6%; Score 31; DB 1; Length 267;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPT 6
DB 232 MYPT 236

RESULT 10
MM07_RAT STANDARD; PRT; 267 AA.
ID MM07_RAT
AC P50280;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE MATRILYSIN PRECURSOR (EC 3.4.24.23) (PUMP-1 PROTEASE) (UTERINE
DE METALLOPROTEINASE) (MATRIX METALLOPROTEINASE-7) (MMP-7) (MATRIN).
GN MMP7 OR MMP-7.

OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Uterus;
RX MEDLINE=95332299; PubMed=7608162;
RA Abramson S.R., Conner G.E., Nagase H., Neuhaus I., Woessner J.F.;
RT "Characterization of rat uterine matrilysin and its cDNA.
RT Relationship to human pump-1 and activation of procollagenases.";
RL J. Biol. Chem. 270:16016-16022(1995).

CC -!- FUNCTION: DEGRADES CASEIN, GELATINS OF TYPES I, III, IV, AND V,
CC AND FIBRONECTIN. ACTIVATES PROCOLLAGENASE (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: CLEAVAGE OF 14-ALA-|-LEU-15 AND 16-TYR-|-
CC LEU-17 IN B CHAIN OF INSULIN. NO ACTION ON COLLAGEN TYPES I, II,
CC IV, V. CLEAVES GELATIN CHAIN ALPHA-2(I) > ALPHA-1(I).
CC -!- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC
CC METALLOPROTEASE) ALSO KNOWN AS MATRININ SUBFAMILY.

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

CC EMBL; L24374; AAA99432.1; -;
CC HSSP; P09237; IMR.
CC MEROPS; M10.008; -;
CC InterPro; IPR000130; -;
CC InterPro; IPR001818; -;
CC PRINTS; PF00413; Peptidase_M10; 1.
CC PROSITE; PS00138; MATRININ.
CC PROSITE; PS00142; ZINC_PROTEASE; 1.
CC PROSITE; PS00546; CYSTEINE_SWITCH; 1.
CC Hydrolase; Metalloprotease; zinc; zymogen; Calcium;
KW Collagen degradation; Extracellular matrix; Signal.

FT SIGNAL 1 20
FT PROPEP 21 97
FT CHAIN 98 267
FT SITE 90 90
FT METAL 217 217
FT ACT_SITE 218 218
FT METAL 221 221
FT METAL 227 227
SQ SEQUENCE 267 AA; 29885 MW; EBA3C3D95274C7B CRC64;
BY SIMILARITY.
ACTIVATION PEPTIDE (BY SIMILARITY).
MATRILYSIN.
CYSTEINE SWITCH (POTENTIAL).
ZINC (CATALYTIC) (BY SIMILARITY).
BY SIMILARITY.
ZINC (CATALYTIC) (BY SIMILARITY).
ZINC (CATALYTIC) (BY SIMILARITY).
ZINC (CATALYTIC) (BY SIMILARITY).

Query Match 75.6%; Score 31; DB 1; Length 267;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPT 6
DB 235 MYPT 239

RESULT 11
SC14_CANAL STANDARD; PRT; 301 AA.
ID SC14_CANAL
AC P46250;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE SEC14 CYTOSOLIC FACTOR (PHOSPHATIDYLINOSITOL/PHOSPHATIDYLCHOLINE
DE TRANSFER PROTEIN) (PI/PC TP).
GN SEC14.

OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5476;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 64385 / 1001.
RX MEDLINE=97051600; PubMed=8896277;
RA Monteguliva L., Sanchez M., Pita J., Gil C., Nombela C.;
RT "Cloning of Candida albicans SEC14 gene homologue coding for a
RT putative essential function.";
RL Yeast 12:1097-1105(1996).

CC -!- FUNCTION: REQUIRED FOR TRANSPORT OF SECRETORY PROTEINS FROM THE
CC GOLGI COMPLEX. CATALYZES THE TRANSFER OF PHOSPHATIDYLINOSITOL AND
CC PHOSPHATIDYLCHOLINE BETWEEN THE MEMBRANES IN VITRO (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: ASSOCIATED WITH THE GOLGI COMPLEX AS A
CC PERIPHERAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE SEC14 CYTOSOLIC FACTOR FAMILY.

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

CC EMBL; X81937; CAA57490.1; -;
CC EMBL; U61975; AAB41491.1; -;
CC HSSP; P24280; IAU.
CC InterPro; IPR001251; -;
CC Pfam; PF00650; CRAL_TRIO; 1.
CC Transport; Protein transport; Golgi stack.
KW SEQUENCE 301 AA; 34709 MW; 1F194A8EA8B525BE CRC64;

Query Match 75.6%; Score 31; DB 1; Length 301;
 Best Local Similarity 100.0%; Pred. No. 36;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPT 6
 |||||
 Db 106 MYPT 110

RESULT 12
 MM12_MOUSE STANDARD; PRT; 462 AA.
 AC P34960;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE MACROPHAGE METALLOELASTASE PRECURSOR (EC 3.4.24.65) (MME) (MATRIX
 DE METALLOPROTEINASE-12) (MMP-12).
 DE MMP12 OR MMEL OR MME.
 Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 99-125.
 RX TISSUE=Macrophage;
 RX MEDLINE=92165826; PubMed=1537850;
 RA Shapiro S.D., Griffin G.L., Gilbert D.J., Jenkins N.A.,
 RA Copeland N.G., Welgus H.G., Senior R.M., Ley T.J.;
 RT "Molecular cloning, chromosomal localization, and bacterial
 expression of a murine macrophage metalloelastase.";
 RL J. Biol. Chem. 267:4664-4671(1992).
 CC -1- FUNCTION: MAY BE INVOLVED IN TISSUE INJURY AND REMODELING. HAS
 CC SIGNIFICANT ELASTOLYTIC ACTIVITY.
 CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF SOLUBLE AND INSOLUBLE ELASTIN.
 CC SPECIFIC CLEAVAGES ARE ALSO PRODUCED AT 14-ALA-1-LEU-15 AND 16-
 CC TYR-1-LEU-17 IN THE B CHAIN OF INSULIN.
 CC -1- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY.
 CC -1- SIMILARITY: CONTAINS 1 HEMOPEXIN-LIKE DOMAIN.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC
 CC METALLOPROTEASE) ALSO KNOWN AS MATRININ SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M82831; AAA39526.1; -.
 CC PIR; A42401; A42401.
 CC HSSP; P03956; ICGL.
 CC MEROPS; M10.009; -.
 CC MGD; MGI:97005; Mmp12.
 CC InterPro; IPR000130; -.
 CC InterPro; IPR000585; -.
 CC InterPro; IPR001818; -.
 CC Pfam; PF00413; Peptidase_M10; 1.
 CC Pfam; PF00045; hemopexin; 4.
 CC PRINTS; PR00138; MATRININ.
 CC PROSITE; PS00024; HEMOPEXIN; 1.
 CC PROSITE; PS00142; ZINC_PROTEASE; 1.
 CC PROSITE; PS00546; CYSTEINE_SWITCH; 1.
 CC Hydrolase; Metalloprotease; Glycoprotein; Zinc; Zymogen; Calcium;
 KW Extracellular matrix; Signal.
 FT SIGNAL 1 17 PROBABLE.
 FT PROPEP 18 98 ACTIVATION PEPTIDE.
 FT CHAIN 99 462 MACROPHAGE METALLOELASTASE.
 FT DOMAIN 272 462 HEMOPEXIN-LIKE.
 FT SITE 85 85 CYSTEINE SWITCH (BY SIMILARITY).

FT METAL 211 211 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACT_SITE 212 212 BY SIMILARITY.
 FT METAL 215 215 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 221 221 ZINC (CATALYTIC) (BY SIMILARITY).
 FT CARBOHYD 21 21 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 74 74 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 310 310 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT DISULFID 275 462 BY SIMILARITY.
 SQ SEQUENCE 462 AA; 53841 MW; BB9625906FIDBEDF CRC64;

Query Match 75.6%; Score 31; DB 1; Length 462;
 Best Local Similarity 100.0%; Pred. No. 55;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPT 6
 |||||
 Db 229 MYPT 233

RESULT 13
 MM12_RAT STANDARD; PRT; 465 AA.
 ID MM12_RAT
 AC Q63341;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE MACROPHAGE METALLOELASTASE PRECURSOR (EC 3.4.24.65) (MME) (MATRIX
 DE METALLOPROTEINASE-12) (MMP-12).
 GN MMP12 OR MMEL.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY;
 RA Cossins J., Clements J., Catlin G.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: MAY BE INVOLVED IN TISSUE INJURY AND REMODELING. HAS
 CC SIGNIFICANT ELASTOLYTIC ACTIVITY (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF SOLUBLE AND INSOLUBLE ELASTIN.
 CC SPECIFIC CLEAVAGES ARE ALSO PRODUCED AT 14-ALA-1-LEU-15 AND 16-
 CC TYR-1-LEU-17 IN THE B CHAIN OF INSULIN.
 CC -1- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY (BY SIMILARITY).
 CC -1- SIMILARITY: CONTAINS 1 HEMOPEXIN-LIKE DOMAIN.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC
 CC METALLOPROTEASE) ALSO KNOWN AS MATRININ SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X98517; CAA67142.1; -.
 CC HSSP; P03956; ICGL.
 CC MEROPS; M10.009; -.
 CC InterPro; IPR000130; -.
 CC InterPro; IPR000585; -.
 CC InterPro; IPR001818; -.
 CC Pfam; PF00413; Peptidase_M10; 1.
 CC Pfam; PF00045; hemopexin; 4.
 CC PRINTS; PR00138; MATRININ.
 CC PROSITE; PS00024; HEMOPEXIN; 1.
 CC PROSITE; PS00142; ZINC_PROTEASE; 1.
 CC PROSITE; PS00546; CYSTEINE_SWITCH; 1.
 CC Hydrolase; Metalloprotease; Glycoprotein; Zinc; Zymogen; Calcium;
 KW Extracellular matrix; Signal.
 FT SIGNAL 1 21 PROBABLE.
 FT PROPEP 22 101 ACTIVATION PEPTIDE (BY SIMILARITY).

```

FT CHAIN 102 465 MACROPHAGE METALLOELASTASE.
FT DOMAIN 275 465 HEMOPEXIN-LIKE.
FT SITE 88 88 CYSTEINE SWITCH (BY SIMILARITY).
FT METAL 214 214 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 215 215 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 218 218 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 224 224 ZINC (CATALYTIC) (BY SIMILARITY).
FT CARBOHYD 313 313 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 278 465 BY SIMILARITY.
SQ SEQUENCE 465 AA; 53738 MW; E779B6014EC6FF68 CRC64;

Query Match 75.6%; Score 31; DB 1; Length 465;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
DB 232 MYPTY 236

RESULT 14
MM18_XENLA STANDARD; PRT; 467 AA.
AC O13065;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE MATRIX METALLOPROTEINASE-18 PRECURSOR (EC 3.4.24.-) (MMP-18)
DE (COLLAGENASE-4) (COLLAGENASE 4) (XCOLA4).
GN MMP18 OR COL4.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Intestine;
RX MEDLINE=97035976; PubMed=8898355;
RA Stow N.A., Baizon D.D., Li J., Sedgwick T., Liang V.C.-T.,
RA Sang Q.A., Shi Y.-B.;
RT laevis: possible roles during frog development.;
RL MOL. Biol. Cell 7:1471-1483(1996).
CC -!- FUNCTION: CLEAVES COLLAGEN TYPE I. MAY PLAY A ROLE IN LARVAL
CC TISSUE DEGENERATION AND ADULT ORGANOGENESIS DURING AMPHIBIAN
CC METAMORPHOSIS. MAY BE INVOLVED IN TAIL RESORPTION.
CC -!- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY (BY SIMILARITY).
CC -!- ENZYME REGULATION: UPREGULATED IN THE TAIL BY THYROID HORMONE.
CC -!- SUBCELLULAR LOCATION: SECRETED. ASSOCIATED WITH THE EXTRACELLULAR
CC MATRIX (BY SIMILARITY).
CC -!- TISSUE SPECIFICITY: EXPRESSED ONLY TRANSIENTLY IN WHOLE ANIMAL, AT
CC TIME WHEN TADPOLE FEEDING BEGINS.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED AT HIGH LEVEL AS THE TADPOLE TAIL
CC RESORBS AND DURING HINDLIMB MORPHOGENESIS AND INTESTINAL
CC REMODELING.
CC -!- SIMILARITY: CONTAINS 1 HEMOPEXIN-LIKE DOMAIN.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC
CC METALLOPROTEASE) ALSO KNOWN AS MATRIXIN SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L76275; AAB53148.1;
CC HSSP; P03956; ACGL.
CC MEROPS; M10.018;
DR InterPro; IPR000130;

```

```

DR InterPro; IPR000585;
DR InterPro; IPR001818;
DR Pfam; PF00045; hemopexin; 4.
DR Pfam; PF00413; Peptidase_M10; 1.
DR PRINTS; PRO0138; MATRIXIN.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
DR PROSITE; PS00546; CYSTEINE_SWITCH; FALSE_NEG.
DR PROSITE; PS00024; HEMOPEXIN; FALSE_NEG.
KW Hydrolase; Metalloprotease; Zinc; Calcium; Zymogen; Signal;
KW Collagen degradation; Extracellular matrix.
FT SIGNAL 1 17 POTENTIAL.
FT PROPEP 18 99 BY SIMILARITY.
FT CHAIN 100 467 MATRIX METALLOPROTEINASE-18.
FT DOMAIN 277 467 HEMOPEXIN-LIKE.
FT SITE 92 92 CYSTEINE SWITCH (POTENTIAL).
FT METAL 218 218 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 219 219 BY SIMILARITY.
FT METAL 222 222 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 228 228 ZINC (CATALYTIC) (BY SIMILARITY).
FT DISULFID 280 467 POTENTIAL.
SQ SEQUENCE 467 AA; 52812 MW; 4623F6CEF3454051 CRC64;

Query Match 75.6%; Score 31; DB 1; Length 467;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
DB 236 MYPTY 240

RESULT 15
MM20_MOUSE STANDARD; PRT; 482 AA.
AC P57748;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE MATRIX METALLOPROTEINASE-20 PRECURSOR (EC 3.4.24.-) (MMP-20) (ENAMEL
DE METALLOPROTEINASE) (ENAMELISIN).
GN MMP20.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SVJ;
RX MEDLINE=20079167; PubMed=10610728;
RA Caterina J., Shi J., Krakora S., Bartlett J.D., Engler J.A.,
RA Kozak C.A., Birkedal-Hansen H.;
RT "Isolation, characterization, and chromosomal location of the mouse
RT enamelysin gene.";
RL Genomics 62:308-311(1999).
CC -!- FUNCTION: DEGRADATES AMELOGENIN, THE MAJOR PROTEIN COMPONENT OF THE
CC ENAMEL MATRIX AND TWO OF THE MACROMOLECULES CHARACTERISING THE
CC CARTILAGE EXTRACELLULAR MATRIX: AGGREGAN AND THE CARTILAGE
CC OLIGOMERIC MATRIX PROTEIN (COMP). MAY PLAY A CENTRAL ROLE IN TOOTH
CC ENAMEL FORMATION (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: CLEAVES AGGREGAN AT THE 360-ASN-1-PHE-361 SITE
CC (BY SIMILARITY).
CC -!- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: SECRETED. ASSOCIATED WITH THE EXTRACELLULAR
CC MATRIX (BY SIMILARITY).
CC -!- TISSUE SPECIFICITY: EXPRESSED SPECIFICALLY IN THE ENAMEL ORGAN.
CC -!- PFM: AUTOACTIVATES AT LEAST AT THE 106-ASN-1-TYR-107 SITE (BY
CC SIMILARITY).
CC -!- SIMILARITY: CONTAINS 1 HEMOPEXIN-LIKE DOMAIN.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC
CC METALLOPROTEASE) ALSO KNOWN AS MATRIXIN SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; AF156956; AAF28472.1; -
DR EMBL; AF156947; AAF28472.1; JOINED.
DR EMBL; AF156948; AAF28472.1; JOINED.
DR EMBL; AF156949; AAF28472.1; JOINED.
DR EMBL; AF156950; AAF28472.1; JOINED.
DR EMBL; AF156951; AAF28472.1; JOINED.
DR EMBL; AF156952; AAF28472.1; JOINED.
DR EMBL; AF156953; AAF28472.1; JOINED.
DR EMBL; AF156954; AAF28472.1; JOINED.
DR EMBL; AF156955; AAF28472.1; JOINED.
DR EMBL; AF155933; AAF28470.1; -
DR MEROPS; M10.019; -
DR MGD; MGI:1353466; Mmp20.
DR PROSITE; PS00546; CYSTEINE_SWITCH; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
DR PROSITE; PS00024; HEMOPEXIN; FALSE_NEG.
KW Hydrolase; Metalloprotease; Zinc; Calcium; Zymogen; Signal;
KW Extracellular matrix.
FT SIGNAL 1 21 POTENTIAL.
FT PROPEP 22 106 BY SIMILARITY.
FT CHAIN 107 482 MATRIX METALLOPROTEINASE-20.
FT DOMAIN 292 482 HEMOPEXIN-LIKE.
FT SITE 99 99 CYSTEINE SWITCH (POTENTIAL).
FT METAL 225 225 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 226 226 BY SIMILARITY.
FT METAL 229 229 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 235 235 ZINC (CATALYTIC) (BY SIMILARITY).
FT DISULFID 295 482 BY SIMILARITY.
SQ SEQUENCE 482 AA; 54373 MW; 366149FAF2BDC8BB CRC64;

Query Match 75.6%; Score 31; DB 1; Length 482;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 MYPTY 6
DB 243 MYPTY 247

rch completed: September 4, 2001, 15:54:09
time: 306 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2001, 15:53:36 ; Search time 80.44 Seconds
(without alignments)
13.158 Million cell updates/sec

Title: US-09-630-345-4
Perfect score: 41
Sequence: 1 MYPTYLK 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

total number of hits satisfying chosen parameters: 425036

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_16:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_unclassified:*
- 13: sp_vertebrate:*
- 14: sp_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35	85.4	1088	4 Q9H2T7	Q9h2t7 homo sapien
2	33	80.5	269	1 Q27159	Q27159 methanobact
3	33	80.5	1069	10 Q9L7Z8	Q9l7z8 arabidopsis
4	32	78.0	252	2 Q9L875	Q9l875 haemophilus
5	32	78.0	553	11 Q70577	Q70577 mus musculu
6	32	78.0	555	11 Q9R0W2	Q9r0w2 rattus norv
7	32	78.0	593	11 P97558	P97558 rattus norv
8	32	78.0	593	11 P70485	P70485 rattus norv
9	32	78.0	986	4 Q94858	Q94858 homo sapien
10	32	78.0	1346	4 Q9UI23	Q9ui23 homo sapien
11	31	75.6	40	14 Q9IWZ8	Q9iwz8 human coxa
12	31	75.6	40	14 Q9IWZ7	Q9iwz7 human coxa
13	31	75.6	50	14 Q66796	Q66796 human echov
14	31	75.6	53	6 Q9XS83	Q9xs83 equus cabal
15	31	75.6	198	14 Q9Q907	Q9q907 shope fibro
16	31	75.6	261	4 Q9NR81	Q9nr81 homo sapien
17	31	75.6	261	4 Q9NR87	Q9nr87 homo sapien
18	31	75.6	261	4 Q9G2S2	Q9gzs2 homo sapien
19	31	75.6	262	5 Q19918	Q19918 caenorhabdi

20	31	75.6	267	6 Q9TV55	Q9tv55 sus scrofa
21	31	75.6	291	5 Q94467	Q94467 dictyostell
22	31	75.6	307	5 Q9VJ65	Q9vj65 drosophila
23	31	75.6	332	2 Q51888	Q51888 prevotella
24	31	75.6	378	10 Q9SHH5	Q9shh5 arabidopsis
25	31	75.6	385	10 Q9SKA2	Q9ska2 arabidopsis
26	31	75.6	420	4 Q9H4D2	Q9h4d2 homo sapien
27	31	75.6	445	5 Q9VFA2	Q9vfa2 drosophila
28	31	75.6	483	13 Q98857	Q98857 cynops pyrr
29	31	75.6	497	14 Q96640	Q96640 bovine papi
30	31	75.6	502	5 Q18426	Q18426 caenorhabdi
31	31	75.6	516	5 Q18198	Q18198 caenorhabdi
32	31	75.6	565	1 Q29350	Q29350 archaeoglob
33	31	75.6	585	5 Q9TXN1	Q9txnl caenorhabdi
34	31	75.6	769	11 Q9ESN3	Q9esn3 mus musculu
35	31	75.6	2185	14 Q88445	Q88445 swine vesic
36	30	73.2	137	14 Q9J919	Q9j919 human coxa
37	30	73.2	141	5 Q02173	Q02173 caenorhabdi
38	30	73.2	152	2 Q9I401	Q9i401 pseudomonas
39	30	73.2	155	14 Q98534	Q98534 paramedum
40	30	73.2	158	5 Q9VUZ1	Q9vuz1 drosophila
41	30	73.2	188	5 Q9V5Y8	Q9v5y8 drosophila
42	30	73.2	238	14 Q9QB94	Q9qb94 yaba monkey
43	30	73.2	248	2 Q48871	Q48871 lactobacill
44	30	73.2	255	14 Q9DHM6	Q9dhm6 yaba-like d
45	30	73.2	262	2 Q9PIA0	Q9pia0 campylobact

ALIGNMENTS

RESULT 1.

Q9H2T7 ID Q9H2T7 PRELIMINARY; PRT; 1088 AA.
AC Q9H2T7;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE RANBP17.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20568271; PubMed=11024021;
RA Kutay U., Hartmann E., Treichel N., Calado A., Carmo-Fonseca M.,
RA Prehn S., Kraft R., Gorlich D., Bischoff F.R.;
RT "Identification of Two Novel RanGTP-binding Proteins Belonging to the
RT Importin beta Superfamily.";
RL J. Biol. Chem. 275:40163-40168(2000).
DR EMBL; AF222747; AAG4255.1;
SQ SEQUENCE 1088 AA; 124374 MW; A95D5599388AEFB CRC64;

Query Match 85.4%; Score 35; DB 4; Length 1088;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 MYPTYL 7
|||||
Db 744 MYPTYL 749

RESULT 2
Q27159 ID Q27159 PRELIMINARY; PRT; 269 AA.
AC Q27159;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE HYPOTHETICAL 31.5 KDA PROTEIN.
GN MTH1087.

OS Methanobacterium thermoautotrophicum.
 OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
 OC Methanobacterium.
 OX NCBI_TaxID=145262;
 RN (1)

RP SEQUENCE FROM N.A.
 RC STRAIN-DELTA H;
 RX MEDLINE=98037514; PubMed=9371463;
 RA Smith D.R., Doucette-Stamm L.A., DeLoughery C., Lee H.-M., DuBois J.,
 RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
 RA Harrison D., Hoang L., Keagle P., Lamm W., Othier B., Qiu D.,
 RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
 RA Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
 RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
 RA Daniels C.J., Mao J.-I., Rice P., Nolling J., Reeve J.N.;
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum
 delCah: functional analysis and comparative genomics.";
 RL J. Bacteriol. 179:7135-7155(1997).
 DR EMBL; AE000879; AAB85576.1; -;
 DR Hypothetical protein.
 SEQUENCE 269 AA; 31513 MW; 4B7D64EBDC6B8EA7 CRC64;

Query Match 80.5%; Score 33; DB 1; Length 269;
 Best Local Similarity 83.3%; Pred. No. 52;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 YPTYLK 8
 |||||
 DB 175 YPTYIK 180

RESULT 3
 ID Q9LTZ8 PRELIMINARY; PRT; 1069 AA.
 AC Q9LTZ8;
 DT 01-OCT-2000 (TRENBLrel. 15, Created)
 DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
 DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)
 DE SIMILARITY TO NEGATIVE REGULATOR OF VESICLE FORMATION.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-COLUMBIA;
 RX Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;
 Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN-COLUMBIA;
 RX MEDLINE=20277480; PubMed=10819329;
 RA Nakamura Y.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
 features of the regions of 4,504,864 bp covered by sixty P1 and TAC
 clones.";
 RL DNA Res. 7:131-135(2000).
 DR EMBL; AB024028; BAA95706.2; -;
 DR EMBL; AP000361; BAA95706.2; JOINED.
 DR InterPro; IPR000379; -;
 DR InterPro; IPR000734; -;
 DR PROSITE; PS00120; LIPASE_SER; UNKNOWN_1.
 SQ SEQUENCE 1069 AA; 118974 MW; 0BA3545568EC6B44 CRC64;

Query Match 80.5%; Score 33; DB 10; Length 1069;
 Best Local Similarity 83.3%; Pred. No. 2.3e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYL 7
 |||||

DB 71 MYPTYL 76

RESULT 4
 ID Q9L875 PRELIMINARY; PRT; 252 AA.
 AC Q9L875;
 DT 01-OCT-2000 (TRENBLrel. 15, Created)
 DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
 DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)
 DE BETA 1-4 GLUCOSYLTRANSFERASE LGTF.
 GN LGTF.
 OS Haemophilus ducreyi.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Haemophilus.
 OX NCBI_TaxID=730;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=35000;
 RX MEDLINE=20278119; PubMed=10816485;
 RA Filiatrault M.J., Gibson B.W., Schilling B., Sun S., Munson R.S. Jr.,
 RA Campagnari A.A.;
 RT "Construction and Characterization of Haemophilus ducreyi
 RT Lipooligosaccharide (LOS) Mutants Defective in Expression of
 RT Heptosyltransferase III and beta1,4-Glucosyltransferase;
 RT Identification of LOS Glycoforms Containing Lactosamine Repeats.";
 RL Infect. Immun. 68:3352-3361(2000).
 DR EMBL; AF215936; AAF72876.1; -;
 DR InterPro; IPR001173; -;
 DR Pfam; PF00535; Glycos_transf_2; 1.
 DR Transferase_KW
 SQ SEQUENCE 252 AA; 28736 MW; 20297CB79C070AD4 CRC64;

Query Match 78.0%; Score 32; DB 2; Length 252;
 Best Local Similarity 83.3%; Pred. No. 78;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYL 7
 :|||||
 DB 134 LYPTYL 139

RESULT 5
 ID 070577 PRELIMINARY; PRT; 553 AA.
 AC 070577;
 DT 01-AUG-1998 (TRENBLrel. 07, Created)
 DT 01-AUG-1998 (TRENBLrel. 07, Last sequence update)
 DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)
 DE ORGANIC CATION TRANSPORTER 2.
 GN SLC22A2 OR OCT2
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=129; TISSUE=LIVER;
 RA Mooslehner K.A., Allen N.D.;
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -!- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
 DR EMBL; AJ006036; CAA06827.1; -;
 DR MGD; MGI:1335072; SLC22a2.
 DR InterPro; IPR001066; -;
 DR Pfam; PF00083; sugar_tr; 1.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_2.
 KW Transmembrane.
 SQ SEQUENCE 553 AA; 61830 MW; 1E9744F0D5415483 CRC64;

Query Match 78.0%; Score 32; DB 11; Length 553;

Best Local Similarity 57.1%; Pred. No. 1.8e+02; Indels 0; Gaps 0;
Matches 4; Conservative 3; Mismatches 0;

QY 2 MYPTYLK 8
:||||:
Db 457 LYPTYR 463

RESULT 6
Q9ROW2 PRELIMINARY; PRT; 555 AA.
AC Q9ROW2
DT 01-FEB-1997 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE ORGANIC CATION TRANSPORTER OCT2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
[1]
SEQUENCE FROM N.A.
RC TISSUE=KIDNEY;
RX MEDLINE=93316020; PubMed=10385678;
RA Gruendemann D., Liebich G., Kiefer N., Koester S., Schoemig E.;
RT "Selective substrates for non-neuronal monoamine transporters.";
RL Mol. Pharmacol. 56:1-10(1999).
DR EMBL; Y13154; CAB52215.1; -;
DR InterPro; IPR001066; -;
DR Pfam; PF00083; sugar_tr; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_2.
SQ SEQUENCE 555 AA; 62342 MW; 29521969AEIAC206 CRC64;

Query Match 78.0%; Score 32; DB 11; Length 555;
Best Local Similarity 57.1%; Pred. No. 1.8e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
:||||:
Db 457 LYPTYR 463

RESULT 7
P97558 PRELIMINARY; PRT; 593 AA.
AC P97558
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE ORGANIC CATION TRANSPORTER.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
[1]
SEQUENCE FROM N.A.
RC TISSUE=KIDNEY;
RA Gorboulev V.G., Koepsell H.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
DR EMBL; X98334; CAA66979.1; -;
DR InterPro; IPR001066; -;
DR Pfam; PF00083; sugar_tr; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_2.
KW Ion transport; Transmembrane.
SQ SEQUENCE 593 AA; 66100 MW; 36C3E1B5DC057790 CRC64;

Query Match 78.0%; Score 32; DB 11; Length 593;
Best Local Similarity 57.1%; Pred. No. 2e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
:||||:
Db 457 LYPTYR 463

RESULT 8
P70485 PRELIMINARY; PRT; 593 AA.
AC P70485
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE ORGANIC CATION TRANSPORTER OCT2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
[1]
SEQUENCE FROM N.A.
RP STRAIN=SPRAGUE-DAWLEY; TISSUE=KIDNEY;
RC MEDLINE=96295517; PubMed=8702418;
RA Okuda M., Saito H., Urakami Y., Takano M., Inui K.;
RT "cDNA cloning and functional expression of a novel rat kidney organic
RT cation transporter, OCT2.";
RL Biochem. Biophys. Res. Commun. 224:500-507(1996).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
DR EMBL; D83044; BAA11754.1; -;
DR InterPro; IPR001066; -;
DR Pfam; PF00083; sugar_tr; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_2.
KW Transmembrane.
SQ SEQUENCE 593 AA; 66080 MW; 36C1044E0C04B443 CRC64;

Query Match 78.0%; Score 32; DB 11; Length 593;
Best Local Similarity 57.1%; Pred. No. 2e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
:||||:
Db 457 LYPTYR 463

RESULT 9
O94858 PRELIMINARY; PRT; 986 AA.
ID O94858
AC O94858
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE KIAA0758 PROTEIN (FRAGMENT).
GN KIAA0758.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
RP TISSUE=BRAIN;
RC MEDLINE=99087487; PubMed=9872452;
RX Nagase T., Ishikawa K., Suyama M., Kikuno R., Miyajima N., Tanaka A.,
RA Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XI.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 5:277-286(1998).
DR EMBL; AB018301; BAA34478.1; -;
DR InterPro; IPR000203; -;
DR Pfam; PF00083; -;
DR Pfam; PF00002; 7tm_2; 1.
DR Pfam; PF01825; GPS; 1.

DR PRINTS: PR00249; GPCRSECRETIN.
 DR SMART: SM00303; GFS; 1.
 FT NON_TER 1
 SQ SEQUENCE 986 AA; 108720 MW; EF4875B03BAA0E7D CRC64;

Query Match 78.0%; Score 32; DB 4; Length 986;
 Best Local Similarity 85.7%; Pred. No. 3.4e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
 Db 384 MLPTYLK 390

RESULT 10
 ID Q9UI23 PRELIMINARY; PRT; 1346 AA.

AC Q9UI23; 2000 (TREMELREL. 13, Created)
 DT 01-MAY-2000 (TREMELREL. 13, Last sequence update)
 DE 01-MAR-2001 (TREMELREL. 16, Last annotation update)
 DE DJ369012.1 (KIAA0758 PROTEIN).
 GN DJ369012.1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Senra H.;
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX

CC EMBL: AL096772; CAB61578.1; -
 DR InterPro: IPR000082; -
 DR InterPro: IPR00203; -
 DR InterPro: IPR000361; -
 DR InterPro: IPR000832; -
 DR InterPro: IPR003006; -
 DR Pfam: PF00002; 7cm.2; 1.
 DR Pfam: PF00042; 4g; 2.
 DR Pfam: PF01825; GFS; 1.
 DR PRINTS: PR00249; GPCRSECRETIN.
 DR PROSITE: PS0118; EGF_2; UNKNOWN_1.
 DR SMART: SM00240; SEA; 1.
 SQ SEQUENCE 1346 AA; 149456 MW; 72A9D02B08218A60 CRC64;

Query Match 78.0%; Score 32; DB 4; Length 1346;
 Best Local Similarity 85.7%; Pred. No. 4.7e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
 Db 744 MLPTYLK 750

RESULT 11

ID Q9IW28 PRELIMINARY; PRT; 40 AA.
 AC Q9IW28;
 DT 01-OCT-2000 (TREMELREL. 15, Created)
 DT 01-OCT-2000 (TREMELREL. 15, Last sequence update)
 DT 01-MAR-2001 (TREMELREL. 16, Last annotation update)
 DE VP2 PROTEIN (FRAGMENT).
 GN VP2.
 OS Human coxsackievirus B5.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Enterovirus.
 OX NCBI_TaxID=12074;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=R8597;
 RA Casas I., Palacios G.F., Trallero G., Cisterna D., Freire M.C.,
 RT Tenorio A.; Molecular characterization of human enteroviruses in clinical samples
 RT by different RT nested PCR assays and direct sequencing of
 RT amplified products.;
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF252193; AAF71330.1; -
 DR EMBL: AF252193; AAF71330.1; -
 DR InterPro: IPR001676; -
 DR Pfam: PF00073; rhv; 1.
 DR NON_TER 1
 FT NON_TER 40
 SQ SEQUENCE 40 AA; 4388 MW; 7F998A9865EF9F67 CRC64;

Query Match 75.6%; Score 31; DB 14; Length 40;
 Best Local Similarity 71.4%; Pred. No. 17;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
 Db 26 VMPTYLK 32

RESULT 12

ID Q9IWZ7 PRELIMINARY; PRT; 40 AA.
 AC Q9IWZ7;
 DT 01-OCT-2000 (TREMELREL. 15, Created)
 DT 01-OCT-2000 (TREMELREL. 15, Last sequence update)
 DT 01-MAR-2001 (TREMELREL. 16, Last annotation update)
 DE VP2 PROTEIN (FRAGMENT).
 GN VP2.
 OS Human coxsackievirus B5.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Enterovirus.
 OX NCBI_TaxID=12074;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Casas I., Palacios G.F., Trallero G., Cisterna D., Freire M.C.,
 RT Tenorio A.; Molecular characterization of human enteroviruses in clinical samples
 RT by three different RT nested PCR assays and direct sequencing of
 RT amplified products.;
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF252194; AAF71331.1; -
 DR InterPro: IPR001676; -
 DR Pfam: PF00073; rhv; 1.
 DR NON_TER 1
 FT NON_TER 40
 SQ SEQUENCE 40 AA; 4388 MW; 7F998A9865EF9F67 CRC64;

Query Match 75.6%; Score 31; DB 14; Length 40;
 Best Local Similarity 71.4%; Pred. No. 17;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
 Db 26 VMPTYLK 32

RESULT 13

ID Q66796 PRELIMINARY; PRT; 50 AA.
 AC Q66796;
 DT 01-NOV-1996 (TREMELREL. 01, Created)
 DT 01-NOV-1996 (TREMELREL. 01, Last sequence update)
 DT 01-MAR-2001 (TREMELREL. 16, Last annotation update)
 DE CAPSID PROTEIN VP2 (FRAGMENT).
 GN VP2.
 OS Human echovirus 20.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;

OC Enterovirus.

OX NCBI_TaxID=47508;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=JY-1.

RX MEDLINE=96203959; PubMed=8627260;

RA Huttunen P., Santti J., Pulli T., Hyypia T.;

RT "The major echovirus group is genetically coherent and related to

RT coxsackie B viruses.";

RL J. Gen. Virol. 77:715-725(1996).

DR EMBL; X89549; CAA61727.1; -.

DR InterPro: IPR001676; -.

DR Pfam; PF00073; rhv; 1.

FT NON_TER 1

FT NON_TER 50

SQ SEQUENCE 50 AA; 5448 MW; BBFC58B2511FC2E7 CRC64;

Query Match

Best Local Similarity 75.6%; Score 31; DB 14; Length 50;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYLK 8

DB 31 VWPTYLK 37

RESULT 14

Q9XS83

ID Q9XS83 PRELIMINARY; PRT; 53 AA.

AC Q9XS83;

DT 01-NOV-1999 (TrEMBLrel. 12, Created)

DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)

DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)

DE TUBBY PROTEIN (FRAGMENT).

GN TUB.

OS Equus caballus (Horse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.

OX NCBI_TaxID=9796;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=99160488; PubMed=10051323;

RA Caetano A.R., Pomp D., Murray J.D., Bowling A.T.;

RT "Comparative mapping of 18 equine type I genes assigned by somatic

RT cell hybrid analysis.";

RL Mamm. Genome 10:271-276(1999).

DR EMBL; AF097580; AAD25983.1; -.

DR HSSP; P50586; IC82.

DR InterPro: IPR000007; -.

DR Pfam; PF01167; Tub; 1.

FT NON_TER 1

FT NON_TER 53

SQ SEQUENCE 53 AA; 6244 MW; 13A9ED8C69500D3E CRC64;

Query Match

Best Local Similarity 100.0%; Score 31; DB 6; Length 53;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6

DB 16 MYPTY 20

RESULT 15

Q9Q907

ID Q9Q907 PRELIMINARY; PRT; 198 AA.

AC Q9Q907;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)

DE GP064R.

GN S064R.

OS Shope fibroma virus (strain Kasza) (SFV).

OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;

OC Leporipoxvirus.

OX NCBI_TaxID=10272;

RN [1]

RP SEQUENCE FROM N.A.

RX STRAIN=KASZA;

RX MEDLINE=84165064; PubMed=6323741;

RA Delange A.M., Macaulay C., Block W., Mueller T., McFadden G.;

RT "Tumorigenic poxviruses: construction of the composite physical map of

RT the Shope fibroma virus genome.";

RL J. Virol. 50:408-416(1984).

RN [2]

RP SEQUENCE FROM N.A.

RX STRAIN=KASZA;

RX MEDLINE=92074222; PubMed=1660196;

RA Strayer D.S., Jerng H.H., O'Connor K.;

RT "Sequence and analysis of a portion of the genomes of Shope fibroma

RT virus and malignant rabbit fibroma virus that is important for viral

RT replication in lymphocytes.";

RL Virology 185:585-595(1991).

RN [3]

RP SEQUENCE FROM N.A.

RX STRAIN=KASZA;

RX MEDLINE=20032074; PubMed=10562495;

RA Willer D.O., McFadden G., Evans D.H.;

RT "The complete genome sequence of shope (Rabbit) fibroma virus.";

RL Virology 264:319-343(1999).

DR EMBL; AF170722; AAF17946.1; -.

SQ SEQUENCE 198 AA; 23282 MW; 9D181052C1ED42C5 CRC64;

Query Match

Best Local Similarity 75.6%; Score 31; DB 14; Length 198;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYL 7

DB 110 MYPTFL 115

Search completed: September 4, 2001, 15:53:36

Job time: 289 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2001, 15:50:25 ; Search time 77.49 Seconds
(without alignments)
6.259 Million cell updates/sec

Title: US-09-630-345-5
Perfect score: 41
Sequence: 1 XWPYLYK 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_0601.*
1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	97.6	8	19 AAW50937	Vasoactive intesti
2	40	97.6	8	21 AAB08298	Vasoactive intesti
3	40	97.6	8	21 AAB08306	Amino acid sequenc
4	35	85.4	8	21 AAB08304	Amino acid sequenc
5	35	85.4	8	21 AAB08305	Amino acid sequenc
6	35	85.4	450	21 AAB63104	Human secreted pro
7	32	78.0	807	21 AAV57288	Human GPCR protein
8	32	78.0	966	21 AAV40440	Human brain-derive
9	31	75.6	8	20 AAW75457	Mammalian tub prot
10	31	75.6	12	20 AAW97073	Peptidomimetic cap
11	31	75.6	58	20 AAY12222	Human 5' EST secre

12	31	75.6	173	22 AAB74618	Human 72 kDa gelat
13	31	75.6	173	22 AAB68682	Matrilysin catalyt
14	31	75.6	173	19 AAW76253	Human matrilysin-1
15	31	75.6	264	19 AAW76254	Human matrilysin p
16	31	75.6	271	17 AAR96211	Human recombinant
17	31	75.6	297	21 AAB43585	Human cancer assoc
18	31	75.6	325	14 AAR42750	ced-4 gene II prod
19	31	75.6	459	18 AAW36485	Mouse tub Form 1.
20	31	75.6	459	21 AAB26400	Mouse tub Form 1 p
21	31	75.6	460	18 AAW36488	Human TUB Form 6.
22	31	75.6	460	21 AAB26903	Human TUB Form 6 p
23	31	75.6	462	22 AAB49983	Murine macrophage
24	31	75.6	469	21 AAB10655	BPV1 L1 fusion pro
25	31	75.6	469	21 AAB10660	BPV2 L1 fusion pro
26	31	75.6	483	20 AAW93163	Human MMP-20 prote
27	31	75.6	495	6 AAP51099	Sequence of bovine
28	31	75.6	495	20 AAW88482	Bovine papillomavi
29	31	75.6	505	18 AAW36486	Mouse TUB Form II.
30	31	75.6	505	18 AAW10728	Mouse tub gene pro
31	31	75.6	505	19 AAW54367	Mouse tub polypept
32	31	75.6	505	20 AAW75450	Mouse wild type tu
33	31	75.6	505	21 AAB26901	Mouse tub Form II
34	31	75.6	506	18 AAW36497	Human TUB Form 4.
35	31	75.6	506	18 AAW10729	Human tub gene pro
36	31	75.6	506	19 AAW54368	Human tub polypept
37	31	75.6	506	20 AAW75451	Human wild type tu
38	31	75.6	506	21 AAB26910	Human tub form 4 5
39	31	75.6	512	18 AAW36496	Human TUB Form 3.
40	31	75.6	512	21 AAB26912	Human TUB form 3 5
41	31	75.6	518	18 AAW36495	Human TUB Form 2.
42	31	75.6	518	21 AAB26909	Human TUB form 2 5
43	31	75.6	561	18 AAW36489	Human TUB Form 1.
44	31	75.6	561	21 AAB26904	Human TUB Form 1 p
45	30	73.2	73	14 AAR40041	Fillistata peptide

ALIGNMENTS

RESULT 1
AAW50937
ID AAW50937 standard; peptide; 8 AA.
XX
AC AAW50937;
XX
DT 31-JUL-1998 (first entry)
XX
DE Vasoactive intestinal peptide receptor binding inhibitor (VIP2).
XX
KW Vasoactive intestinal peptide; VIP; antagonist; somatostatin; bombesin;
KW Substance P; cancer; inhibition.
XX
OS Synthetic.
XX
PN EP835662-A2.
XX
PD 15-APR-1998.
XX
PF 11-DEC-1996; 96EP-0309012.
XX
PR 08-OCT-1996; 96US-0727679.
PR 16-AUG-1996; 96IN-0001822.
XX
(NAIM-) NAT INST IMMUNOLOGY.
XX
Jaggl M, Mukherjee R;
XX
WPI; 1998-208959/19.
XX
Composition containing analogues of vasoactive intestinal peptide,
XX
somatostatin - bombesin and substance P, for treatment of tumours
XX
and for inhibiting over-expression of these peptide(s)

PS Claim 1; Page 4; 49pp; English.

XX The invention relates to a new composition which comprises: (i) the
 CC somatostatin analogue SOM2 AGCKNFRDWKPTSDC (3-14 disulphide bridge),
 CC and (ii) at least 4 of the peptides: antagonist of vasoactive
 CC intestinal peptide (VIP1); VIP receptor-binding inhibitor (VIP2); VIP
 CC receptor antagonist (VIP3); somatostatin analogue (SOM1); bombesin
 CC antagonist (BOM1) and substance P antagonist (SPI). Also claimed are
 CC more general compositions containing peptide analogues of somatostatin,
 CC VIP, bombesin and substance P. The compositions are used in human or
 CC veterinary medicine: (a) to kill (or inhibit multiplication of) tumour
 CC or cancer cells, particularly for treatment of leukaemia, lymphoma,
 CC adenocarcinoma of stomach, pancreas or prostate, or cancer of lung,
 CC breast, kidney or particularly rectum and colon; and (b) to prevent,
 CC inhibit or modulate over-expression of e.g. VIP. A wide range of cancer
 CC cells express receptors for VIP, somatostatin, bombesin and/or substance
 CC P. The present sequence represents VIP receptor-binding inhibitor
 CC (VIP2).

XX Sequence 8 AA;

Query Match 97.6%; Score 40; DB 19; Length 8;
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 MYPTYLK 8
 DB 2 myptylk 8
 |||||

RESULT 2
 AAB08298
 ID AAB08298 standard; peptide; 8 AA.
 XX
 AC AAB08298;
 XX
 DT 04-DEC-2000 (first entry)
 DE Vasoactive intestinal peptide (VIP) analogue VIP2.
 XX
 KW Vasoactive intestinal peptide; VIP; analogue; somatostatin; SOM1; SOM2;
 KW VIP1; VIP2; VIP3; BOM1; bombesin; SPI; substance P; MuJ-7; tumour growth;
 KW tumour angiogenesis; metastasis; cancer; angiogenesis; adenocarcinoma;
 KW leukaemia; lymphoma.
 XX
 OS Synthetic.
 XX
 WO2000047221-A1.
 PD 17-AUG-2000.
 XX
 PF 11-FEB-2000; 2000WO-US03559.
 XX
 PR 11-FEB-1999; 99US-0248381.
 XX
 PA (NATIM-) NAT INST IMMUNOLOGY.
 PA (DABU-) DABUR RES FOUND.
 PA (CORD/) CORD J I.
 XX
 PI Mukherjee R, Jaggi M, Prasad S, Burman AC, Rajendran P, Mathur A;
 PI Singh AT;
 XX
 WPI: 2000-549083/50.
 XX
 Novel therapeutically active composition comprising at least 5
 PT peptides, useful for treating angiogenesis especially as a result of
 PT adenocarcinomas -
 XX
 Disclosure; Page 8; 42pp; English.
 PS
 The present sequence represents an analogue of vasoactive intestinal
 CC peptide (VIP). The specification describes therapeutically active

CC compositions comprising at least one analogue of somatostatin (chosen
 CC from SOM1 and SOM2), and at least four analogues chosen from VIP1 (a
 CC VIP antagonist), VIP2 (a VIP receptor binding inhibitor), VIP3 (a VIP
 CC receptor antagonist), BOM1 (a bombesin antagonist), and SPI (a substance
 CC P antagonist). The combination of these 7 analogues is known as MuJ-7.
 CC MuJ-7 is used as an anticancer drug to restrict tumour growth and spread
 CC by inhibiting tumour angiogenesis. MuJ-7, in addition, inhibits
 CC metastasis through its antiangiogenic activity in all cancers. The
 CC peptides are useful for the treatment and prevention of angiogenesis,
 CC especially as a result of adenocarcinomas of the colon, breast, lung,
 CC prostate, kidney, leukemias or lymphomas.

XX Sequence 8 AA;

Query Match 97.6%; Score 40; DB 21; Length 8;
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 MYPTYLK 8.
 DB 2 myptylk 8
 |||||

RESULT 3
 AAB08306
 ID AAB08306 standard; peptide; 8 AA.
 XX
 AC AAB08306;
 XX
 DT 04-DEC-2000 (first entry)
 DE Amino acid sequence of an antiangiogenic peptide.
 XX
 KW Vasoactive intestinal peptide; VIP; analogue; somatostatin; SOM1; SOM2;
 KW VIP1; VIP2; VIP3; BOM1; bombesin; SPI; substance P; MuJ-7; tumour growth;
 KW tumour angiogenesis; metastasis; cancer; angiogenesis; adenocarcinoma;
 KW leukaemia; lymphoma.
 XX
 OS Synthetic.
 XX
 Key Location/Qualifiers
 FH Misc-difference 1 /note= "optionally D-form residue"
 FT Misc-difference 3 /note= "optionally D-form residue"
 FT Misc-difference 6 /note= "optionally D-form residue"
 FT Misc-difference 7 /note= "optionally D-form residue"
 FT Misc-difference 7 /note= "optionally D-form residue"
 XX
 WO2000047221-A1.
 PD 17-AUG-2000.
 XX
 PF 11-FEB-2000; 2000WO-US03559.
 XX
 PR 11-FEB-1999; 99US-0248381.
 XX
 PA (NATIM-) NAT INST IMMUNOLOGY.
 PA (DABU-) DABUR RES FOUND.
 PA (CORD/) CORD J I.
 XX
 PI Mukherjee R, Jaggi M, Prasad S, Burman AC, Rajendran P, Mathur A;
 PI Singh AT;
 XX
 WPI: 2000-549083/50.
 XX
 Novel therapeutically active composition comprising at least 5
 PT peptides, useful for treating angiogenesis especially as a result of
 PT adenocarcinomas -
 XX
 Claim 11; Page 31; 42pp; English.
 PS

XX AAB08304-15 represent peptides which have an antiangiogenic effect. The
 CC specification describes therapeutically active compositions comprising
 CC at least one analogue of somatostatin (chosen from SOM1 and SOM2), and
 CC at least four analogues chosen from vasoactive intestinal peptide (VIP)
 CC 1 (a VIP antagonist), VIP2 (a VIP receptor binding inhibitor), VIP3 (a
 CC VIP receptor antagonist), BOM1 (a bombesin antagonist), and SP1 (a
 CC substance P antagonist). The combination of these 7 analogues is known as
 CC MuJ-7. MuJ-7 is used as an anticancer drug to restrict tumour growth and
 CC spread by inhibiting tumour angiogenesis. MuJ-7, in addition, inhibits
 CC metastasis through its antiangiogenic activity in all cancers. The
 CC peptides are useful for the treatment and prevention of angiogenesis,
 CC especially as a result of adenocarcinomas of the colon, breast, lung,
 CC prostate, kidney, leukemias or lymphomas.
 XX SQ Sequence 8 AA;

Query Match 97.6%; Score 40; DB 21; Length 8;
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 MYPTYLK 8
 |||||
 Db 2 myptylk 8

RESULT 4
 AAB08304
 ID AAB08304 standard; peptide; 8 AA.
 XX AC AAB08304;
 XX DT 04-DEC-2000 (first entry)
 XX DE Amino acid sequence of an antiangiogenic peptide.
 XX KW Vasoactive intestinal peptide; VIP; analogue; somatostatin; SOM1; SOM2;
 KW VIP1; VIP2; VIP3; BOM1; bombesin; SP1; substance P; MuJ-7; tumour growth;
 KW tumour angiogenesis; metastasis; cancer; angiogenesis; adenocarcinoma;
 KW leukaemia; lymphoma.
 XX OS Synthetic.

Key Location/Qualifiers
 FH Modified-site 1
 FT /label= Aib
 FT /note= "alpha-aminoisobutyric acid"
 FT Modified-site 7
 FT /label= Aib, Deg
 FT /note= "alpha-aminoisobutyric acid or alpha, alpha
 diethyl glycine"

WO200047221-A1.
 XX 17-AUG-2000.
 XX 11-FEB-2000; 2000WO-US03559.
 XX 11-FEB-1999; 99US-0248381.
 XX (NAIM-) NAT INST IMMUNOLOGY.
 PA (DABU-) DABUR RES FOUND.
 PA (CORD/) CORD J I.
 XX Mukherjee R, Jaggi M, Prasad S, Burman AC, Rajendran P, Mathur A;
 PI Singh AT;
 XX WPI: 2000-549083/50.

XX Novel therapeutically active composition comprising at least 5
 PT peptides, useful for treating angiogenesis especially as a result of
 PT adenocarcinomas -

XX Claim 18; Page 36; 42pp; English.
 XX AAB08304-15 represent peptides which have an antiangiogenic effect. The
 CC specification describes therapeutically active compositions comprising
 CC at least one analogue of somatostatin (chosen from SOM1 and SOM2), and
 CC at least four analogues chosen from vasoactive intestinal peptide (VIP)
 CC 1 (a VIP antagonist), VIP2 (a VIP receptor binding inhibitor), VIP3 (a
 CC VIP receptor antagonist), BOM1 (a bombesin antagonist), and SP1 (a
 CC substance P antagonist). The combination of these 7 analogues is known as
 CC MuJ-7. MuJ-7 is used as an anticancer drug to restrict tumour growth and
 CC spread by inhibiting tumour angiogenesis. MuJ-7, in addition, inhibits
 CC metastasis through its antiangiogenic activity in all cancers. The
 CC peptides are useful for the treatment and prevention of angiogenesis,
 CC especially as a result of adenocarcinomas of the colon, breast, lung,
 CC prostate, kidney, leukemias or lymphomas.
 XX SQ Sequence 8 AA;

Query Match 85.4%; Score 35; DB 21; Length 8;
 Best Local Similarity 85.7%; Pred. No. 3.4e+05;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
 |||||
 Db 2 myptylk 8

RESULT 5
 AAB08305
 ID AAB08305 standard; peptide; 8 AA.
 XX AC AAB08305;
 XX DT 04-DEC-2000 (first entry)

Amino acid sequence of an antiangiogenic peptide.

XX Vasoactive intestinal peptide; VIP; analogue; somatostatin; SOM1; SOM2;
 KW VIP1; VIP2; VIP3; BOM1; bombesin; SP1; substance P; MuJ-7; tumour growth;
 KW tumour angiogenesis; metastasis; cancer; angiogenesis; adenocarcinoma;
 KW leukaemia; lymphoma.

OS Synthetic.
 FH Key Location/Qualifiers
 FT Misc-difference 1
 FT /note= "D-form residue"
 FT Modified-site 7
 FT /label= Aib, Deg
 FT /note= "alpha-aminoisobutyric acid or alpha, alpha
 diethyl glycine"

WO200047221-A1.
 XX 17-AUG-2000.
 XX 11-FEB-2000; 2000WO-US03559.
 XX 11-FEB-1999; 99US-0248381.
 XX (NAIM-) NAT INST IMMUNOLOGY.
 PA (DABU-) DABUR RES FOUND.
 PA (CORD/) CORD J I.

XX Mukherjee R, Jaggi M, Prasad S, Burman AC, Rajendran P, Mathur A;
 PI Singh AT;
 XX WPI: 2000-549083/50.

XX Novel therapeutically active composition comprising at least 5
 PT peptides, useful for treating angiogenesis especially as a result of

PT adenocarcinomas -

PS Claim 11; Page 31; 42pp; English.

XX

CC AA08304-15 represent peptides which have an antiangiogenic effect. The

CC specification describes therapeutically active compositions comprising

CC at least one analogue of somatostatin (chosen from SOM1 and SOM2), and

CC at least four analogues chosen from vasoactive intestinal peptide (VIP)

CC 1 (a VIP antagonist); VIP2 (a VIP receptor binding inhibitor); VIP3 (a

CC VIP receptor antagonist); BOM1 (a bombesin binding inhibitor); VIP3 (a

CC substance P antagonist); BOM1 (a bombesin antagonist); and SPl (a

CC MuJ-7 MuJ-7 antagonist). The combination of these 7 analogues is known as

CC spread by inhibiting tumour angiogenesis. MuJ-7, in addition, inhibits

CC metastasis through its antiangiogenic activity in all cancers. The

CC peptides are useful for the treatment and prevention of angiogenesis,

CC especially as result of adenocarcinomas of the colon, breast, lung,

CC prostate, kidney, leukemias or lymphomas.

XX

SO Sequence 8 AA;

Query Match 85.4%; Score 35; DB 21; Length 8;

Best Local Similarity 85.7%; Pred. No. 3.4e+05;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MYPTYLK 8

DB 2 myptylk 8

|||||

26-MAR-2001 (first entry)

Human secreted protein sequence encoded by gene 18 SEQ ID NO:114.

Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;

antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;

cerebroprotective; neurotropic; neuroprotective; antibacterial; virucide;

fungicide; ophthalmological; vulnary; gene therapy; neoplasia; virucide;

autoimmune disease; rheumatoid arthritis; hyperproliferative disorder;

cardiovascular disorder; cardiac arrest; cerebrovascular disorder;

cerebral ischaemia; angiogenesis; nervous system disorder; infection;

Alzheimer's disease; ocular disorder; corneal infection; wound healing;

skin aging; food additive; preservative.

OS Homo sapiens.

XX WO200061748-A1.

PN 19-OCT-2000.

XX

XX 06-APR-2000; 2000WO-US08982.

XX

XX 09-APR-1999; 99US-0128696.

XX

XX 14-JAN-2000; 2000US-0176069.

XX

XX (HUMA-) HUMAN GENOME SCI INC.

XX

XX Rosen CA, Ruben SM, Komatsoulis G;

XX

XX WPI; 2000-638566/61.

XX

XX New nucleic acid molecules encoding 48 human secreted proteins for

XX diagnosing, preventing, treating or ameliorating medical conditions and

XX used as food additives or preservatives -

XX

PS Disclosure; Page 453-455; 480pp; English.

CC AAF22316 to AAF22363 encode the human secreted proteins given in AAB63049

CC to AAB63096. AAB63097 to AAB63132 represent more human secreted proteins

CC and polypeptides homologous to them. Human secreted proteins have

CC activities based on the tissues and cells the genes are expressed in.

CC Examples of activities include: immunosuppressive; antiarthritic;

CC antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;

CC cerebroprotective; neurotropic; neuroprotective; antibacterial; virucide;

CC fungicide; ophthalmological; and vulnary. The polynucleotides and

CC proteins can be used to prevent, treat or ameliorate a medical condition

CC in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or

CC sheep. They are also used in diagnosing a pathological condition or

CC susceptibility to a pathological condition. Disorders which are diagnosed

CC or treated include autoimmune diseases e.g. rheumatoid arthritis,

CC hyperproliferative disorders e.g. neoplasms of the breast or liver,

CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders

CC e.g. cerebral ischaemia, infectious diseases caused by bacteria, viruses and fungi and

CC Alzheimer's disease. Infections caused by bacteria, viruses and fungi and

CC ocular disorders e.g. corneal infection. The polypeptides can also be

CC used to aid wound healing and epithelial cell proliferation, to prevent

CC skin aging due to sunburn, to maintain organs before transplantation, for

CC chemotaxis. The polypeptides can also be used as a food additive or

CC preservative to increase or decrease storage capabilities, fat content,

CC lipid, protein, carbohydrate, vitamins, minerals, cofactors and other

CC nutritional components. AAF22307 to AAF22315 and AAB63048 represent

CC sequences used in the exemplification of the present invention.

XX

SO Sequence 450 AA;

Query Match 85.4%; Score 35; DB 21; Length 450;

Best Local Similarity 100.0%; Pred. No. 40;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYL 7

DB 106 myptyl 111

|||||

RESULT 7

AA57288

ID AAY57288 standard; Protein; 807 AA.

XX

XX AAY57288;

XX

XX 05-JUN-2000 (first entry)

XX

XX Human GPCR protein (HGRP) sequence (clone ID 3036563).

XX

XX Human; G protein coupled protein receptor; HGRP; cell proliferation;

XX neurological; immune disorder; cytostatic; anti-arteriosclerotic;

XX anti-atherosclerotic; hepatotropic; antinflammatory; virucide; leukemia;

XX immunomodulatory; anemia; asthma; gastrointestinal; anti-epileptic;

XX anti-Alzheimer's; anti-Parkinsonian; gene therapy.

XX

XX Homo sapiens.

XX

XX WO200015793-A2.

XX

XX 23-MAR-2000.

XX

XX 17-SEP-1999; 99WO-US20958.

XX

XX 17-SEP-1998; 98US-0156513.

XX

XX (INCY-) INCYTE PHARM INC.

XX

XX Bandman O, Lal P, Tang YT, Corley NC, Guegler KJ, Gorgone GA;

XX Baughn MR;

XX

XX WPI; 2000-271432/23.

XX

XX N-PSDB; AA290526.

PT Human G protein coupled protein receptor peptides useful for the
 PT prevention, diagnosis and treatment of cell proliferative, neurological
 PT and immune disorders -

XX Claim 1; Page 65-67; 71pp; English.

XX The invention provides human G protein coupled protein receptor (HGPRP)
 CC polypeptides and polynucleotides encoding them. The polypeptides can be
 CC produced by standard recombinant methodology. The polynucleotides and
 CC polypeptides may be used in the prevention, treatment and diagnosis of
 CC diseases associated with their inappropriate expression. Diseases that
 CC can be treated are cell proliferative disorders (e.g. arteriosclerosis,
 CC atherosclerosis and hepatitis), cancers (e.g. leukemia, melanomas and
 CC adenocarcinoma), immune disorders (e.g. anemia, asthma and Crohn's
 CC disease) and neurological disorders (e.g. epilepsy, Alzheimer's disease
 CC and Parkinson's disease). The anti-HGPRP antibodies may also be used as
 CC diagnostic agents for detecting the presence of HGPRP polypeptides in
 CC samples (e.g. by enzyme linked immunosorbent assay (ELISA)). Sequences
 CC AAY57283-288 represent the HGPRP polypeptides.

XX Sequence 807 AA;

Query Match 78.0%; Score 32; DB 21; Length 807;

Best Local Similarity 85.7%; Pred. No. 2.9e+02; Mismatches 1; Indels 0; Gaps 0;

XX 2 MYPTYLK 8

Db 205 mlptylk 211

RESULT 8

AAAY40440

ID AAY40440 standard; Protein; 986 AA.

XX AC AAY40440;

XX 26-MAY-2000 (first entry)

XX Human brain-derived G-protein coupled receptor protein.

XX G-protein coupled receptor protein; human; brain; gene therapy;
 KW genetic disease; screening assay.

XX Homo sapiens.

XX WO200008053-A1.

XX 17-FEB-2000.

XX 05-AUG-1999; 99WO-JP04233.

XX 07-AUG-1998; 98JP-0225059.

XX (TAKE) TAKEDA CHEM IND LTD.

XX (KAZU-) KAZUSA DNA RES INST.

XX Ohara O, Nagase T, Nomura N, Mogi S, Yamamoto K, Kurokawa T;

XX WPI; 2000-195555/17.

XX N-PSDB; AA287930.

XX Human brain-derived G-protein coupled receptor protein and encoding DNA
 PT used in gene therapy, is also useful for e.g. identifying ligands and
 PT raising antibodies and antisera -

XX Claim 1; Fig 1; 80pp; Japanese.

XX This represents a human brain-derived G-protein coupled receptor
 CC protein. The polypeptide can be expressed by standard recombinant
 CC methodology. The novel G-protein coupled receptor protein can be used
 CC for identifying ligands; raising antibodies and antisera; developing

CC receptor-binding assay system; and screening for drug candidates. The
 CC encoding polynucleotide can be used in; probing for diagnostic genes;
 CC constructing PCR primers; making transgenic animals; and in gene therapy.

XX Sequence 986 AA;

Query Match 78.0%; Score 32; DB 21; Length 986;

Best Local Similarity 85.7%; Pred. No. 3.6e+02; Mismatches 1; Indels 0; Gaps 0;

XX 2 MYPTYLK 8

Db 384 mlptylk 390

RESULT 9

AAW75457

ID AAW75457 standard; peptide; 8 AA.

XX AC AAW75457;

XX 27-APR-1999 (first entry)

XX Mammalian tub protein tyrosine phosphorylation site.

XX Human; wild type; tubby; identification; SH2 domain; mammal; obesity;
 KW body weight disorder; cachexia; anorexia.

XX Homo sapiens.

XX Mus sp.

XX US5861239-A.

XX 19-JAN-1999.

XX 02-SEP-1997; 97US-0922267.

XX 02-SEP-1997; 97US-0922267.

XX 12-APR-1996; 96US-0631200.

XX 28-MAR-1997; 97US-0829553.

XX (MILL-) MILLENNIUM PHARM INC.

XX Kapeller R, Kleyn PW, Moore KJ;

XX WPI; 1999-130383/11.

XX Identifying compounds which modulate tub protein activity - by
 PT detecting compounds which alter the interaction of tub protein with
 PT a SH2 containing peptide, used to develop agents for treating e.g.
 PT obesity, cachexia or anorexia

XX Disclosure; Column 13; 95pp; English.

XX This sequence represents a tyrosine phosphorylation site/SH2 docking
 CC domain found in the amino acid sequence of the mouse and human "tub"
 CC proteins (AAW75450 and AAW75451 respectively). The invention relates to
 CC a method for identifying compounds that modulate tub protein activity,
 CC especially its interaction with proteins containing an SH2 domain. The
 CC method can be used for identifying compounds which modulate tub protein
 CC activity for use in the treatment of mammalian body weight disorders
 CC including obesity, cachexia and anorexia.

XX Sequence 8 AA;

Query Match 75.6%; Score 31; DB 20; Length 8;

Best Local Similarity 100.0%; Pred. No. 3.4e+05; Mismatches 0; Indels 0; Gaps 0;

XX 2 MYPTPY 6

XX |||||

KW excess connective tissue removal; dermatological; keloid; scleroderma;
 KW post-operative fibrosis; intervertebral disc injection; fibrotic disease;
 KW hypertrophic scar; wound debridement; post-surgical adhesion;
 KW idiopathic pulmonary fibrosis.

XX Homo sapiens.

OS US6194189-B1.

PN 27-FEB-2001.

PD 16-DEC-1994; 94US-0357820.

XX 16-DEC-1994; 94US-0357820.

PR (UNIW) UNIV WASHINGTON.

XX Senior RM;

PI WPI; 2001-243407/25.

XX New gelatinase truncated mutant useful for treating disorders requiring

PT the removal of excess connective tissue, e.g. keloids, post-operative

PT fibrosis, intervertebral disc injections, hypertrophic scars

XX Example; Column 13-16; l1pp; English.

XX The present invention describes a cDNA sequence which encodes the

CC truncated mutant of the 92 kDa gelatinase having an amino acid sequence

CC as given in AAB74617 consisting of residues 106-216 fused to residues

CC 391-443 of the parent molecule. The truncated mutant has dermatological

CC and vulnary activities. The truncated mutant is useful for treating

CC disorders requiring the removal of excess connective tissue,

CC e.g., keloids, post-operative fibrosis, intervertebral disc injections,

CC hypertrophic scars, wound debridement, post-surgical adhesions and

CC various fibrotic diseases (including scleroderma, idiopathic pulmonary

CC fibrosis). The truncated mutant is catalytically active compared to

CC the full protein. Unlike the full protein, the truncated mutant is

CC essentially inactive against insoluble elastin, and does not require

CC activation to be enzymatically active. The present sequence represents

CC a 72 kDa gelatinase catalytic domain (72 CD) containing protein, which

CC is given in the exemplification of the present invention.

XX Sequence 173 AA;

SQ

Query Match 75.6%; Score 31; DB 22; Length 173;

Best Local Similarity 100.0%; Pred. No. 89;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6

DB 138 mypty 142

RESULT 13

AAB68682

ID AAB68682 standard; Protein; 173 AA.

XX AAB68682;

XX 02-MAY-2001 (first entry)

XX Matrilysin catalytic domain.

XX Gelatinase; excess connective tissue removal; matrilysin;

XX matrix metalloproteinase.

XX Unidentified.

OS US6184021-B1.

XX 06-FEB-2001.

PD

XX

PF 19-MAY-1995; 95US-0444628.

XX 16-DEC-1994; 94US-0357820.

XX (UNIW) UNIV WASHINGTON.

XX Senior RM;

PI WPI; 2001-202001/20.

XX New truncated mutant of 92 kDa gelatinase which is catalytically

XX active, but is inactive against insoluble elastin, useful for treating

XX disorders requiring the removal of excess connective tissues such as

XX keloids

XX Examples; Fig 2; l1pp; English.

XX The present invention relates to a truncated mutant (92 CD) of the 92

XX kilo Dalton (kDa) gelatinase (see AAB58681). The truncated protein is

XX useful for treating disorders requiring the removal of excess connective

XX tissue, e.g. keloids, post-operative fibrosis, intervertebral disc

XX injections, hypertrophic scars, wound debridement, post-surgical

XX adhesions and various fibrotic diseases (scleroderma, idiopathic

XX pulmonary fibrosis). Gelatinase is a matrix metalloproteinase and is also

XX known as gelatinase B and MMP-9. The truncated protein is catalytically

XX active comparable to the full protein but unlike the full protein is

XX essentially inactive against insoluble elastin. The present sequence is

XX the catalytic domain of matrilysin, which was used in a sequence homology

XX alignment with the 92 CD protein of the present invention. Matrilysin is

XX also a matrix metalloproteinase.

XX Sequence 173 AA;

SQ

Query Match 75.6%; Score 31; DB 22; Length 173;

Best Local Similarity 100.0%; Pred. No. 89;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6

DB 138 mypty 142

RESULT 14

AAW76253

ID AAW76253 standard; Protein; 261 AA.

XX AAW76253;

XX 02-DEC-1998 (first entry)

XX Human matrilysin-like protein.

XX Matrilysin; inhibitor; tumour necrosis factor; screening; MMP; infection;

XX matrix metalloproteinase; antagonist; diagnosis; underexpression; cancer;

XX degeneration; extracellular matrix; arthritis; cardiovascular disease;

XX cachexia; multiple sclerosis.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..22 /label= signal

FT Protein 23..261 /label= matrilysin_like_protein

XX WO9831818-A2.

XX 23-JUL-1998.

XX 20-JAN-1998; 98WO-US00783.

XX

```

- PR 01-AUG-1997; 97US-0054541.
PR 21-JAN-1997; 97US-0034205.
PR 13-JUN-1997; 97US-0049607.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Brewer L, Gentz R, Ni J, Rosen CA, Ruben SM;
XX WPI; 1998-414114/35.
XX N-PSDB; AAV61633.
XX
XX Isolated nucleic acid encoding human metallo-protease(s) - used for
XX diagnosis, treatment and prevention of, e.g. cancer, inflammation,
XX neurological disease and infections
XX
XX Claim 20a; Fig 4; 8lpp; English.
XX
XX This sequence represents a novel human matrilysin-like protein which is
XX an inhibitor of the members of the matrix metalloproteinase, MMP,
XX family. This protein can be used in assays to screen for agonists and
XX antagonists and the nucleic acid is used as a probe for gene mapping, in
XX situ hybridisation and detection of corresponding genes in human tissue,
XX and as sources of probes and primers for diagnosis. The protein and its
XX antigenic fragments are used to raise antibodies (Ab) (which can be used
XX for diagnosis). Agonists can be used for in vivo imaging) and to screen
XX conditions associated with underexpression of TNF-alpha. Antagonists are
XX used to treat diseases associated with degeneration of the extracellular
XX matrix (cancer, arthritis, cardiovascular disease, cachexia and multiple
XX sclerosis).
XX
XX Sequence 261 AA;

Query Match 75.6%; Score 31; DB 19; Length 261;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
DB 226 mypty 230
|||||

RESULT 15
AAW76254
ID AAW76254 standard; Protein; 264 AA.
XX
XX AAW76254;
XX
XX 02-DEC-1998 (first entry)
XX Human matrilysin protein.
XX
XX Matrilysin; inhibitor; tumour necrosis factor; screening; MMP; infection;
XX matrix metalloproteinase; antagonist; diagnosis; underexpression; cancer;
XX degeneration; extracellular matrix; arthritis; cardiovascular disease;
XX cachexia; multiple sclerosis.
XX
XX Homo sapiens.
XX
XX WO9831818-A2.
XX
XX 23-JUL-1998.
XX
XX 20-JAN-1998; 98WO-US00783.
XX
XX 01-AUG-1997; 97US-0054541.
XX 21-JAN-1997; 97US-0034205.
XX 13-JUN-1997; 97US-0049607.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Brewer L, Gentz R, Ni J, Rosen CA, Ruben SM;

```

```

XX WPI; 1998-414114/35.
XX
XX Isolated nucleic acid encoding human metallo-protease(s) - used for
XX diagnosis, treatment and prevention of, e.g. cancer, inflammation,
XX neurological disease and infections
XX
XX Disclosure; Fig 5; 8lpp; English.
XX
XX This sequence represents the human matrilysin protein which is an
XX inhibitor of the members of the matrix metalloproteinase, MMP, family.
XX This protein is used in the identification of a novel human
XX matrilysin-like protein which can be used in assays to screen for
XX agonists and antagonists and the nucleic acid is used as a probe for gene
XX mapping, in situ hybridisation and detection of corresponding genes in
XX human tissue, and as sources of probes and primers for diagnosis. The
XX protein and its antigenic fragments are used to raise antibodies (Ab)
XX (which can be used for diagnosis in usual immunoassays or for in vivo
XX imaging) and to screen for (antagonists). Agonists can be used to treat
XX cancer or other conditions associated with underexpression of TNF-alpha.
XX Antagonists are used to treat diseases associated with degeneration of
XX the extracellular matrix (cancer, arthritis, cardiovascular disease,
XX cachexia and multiple sclerosis).
XX
XX Sequence 264 AA;

Query Match 75.6%; Score 31; DB 19; Length 264;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
DB 232 mypty 236
|||||

Search completed: September 4, 2001, 15:50:25
Job time: 163 sec

```

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2001, 15:51:13 ; Search time 40.98 seconds
(without alignments)
4.020 Million cell updates/sec

Title: US-09-630-345-5

Perfect score: 41

Sequence: 1 XMYPTYLK 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents.AA.*

1: /cgn2.6/ptodata/2/1aa/5A-COMB.pep.*

2: /cgn2.6/ptodata/2/1aa/5b-COMB.pep.*

3: /cgn2.6/ptodata/2/1aa/6A-COMB.pep.*

4: /cgn2.6/ptodata/2/1aa/6B-COMB.pep.*

5: /cgn2.6/ptodata/2/1aa/PCTUS-COMB.pep.*

6: /cgn2.6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	97.6	8	1 US-07-620-410-3	Sequence 3, Appli
2	31	75.6	8	2 US-08-922-267A-65	Sequence 65, Appl
3	31	75.6	173	4 US-08-444-628-8	Sequence 8, Appli
4	31	75.6	173	4 US-08-357-820-8	Sequence 8, Appli
5	31	75.6	261	3 US-09-009-156-5	Sequence 5, Appli
6	31	75.6	264	3 US-09-009-156-5	Sequence 6, Appli
7	31	75.6	267	4 US-08-448-489-18	Sequence 18, Appli
8	31	75.6	271	3 US-08-896-062-2	Sequence 2, Appli
9	31	75.6	459	1 US-08-630-592-2	Sequence 2, Appli
10	31	75.6	459	1 US-08-714-991-2	Sequence 2, Appli
11	31	75.6	459	3 US-09-032-365A-2	Sequence 2, Appli
12	31	75.6	460	1 US-08-630-592-7	Sequence 7, Appli
13	31	75.6	460	1 US-08-714-991-7	Sequence 7, Appli
14	31	75.6	460	3 US-09-032-365A-8	Sequence 8, Appli
15	31	75.6	462	4 US-08-068-392-3	Sequence 3, Appli
16	31	75.6	462	4 US-08-396-988-3	Sequence 3, Appli
17	31	75.6	505	1 US-08-631-200-2	Sequence 2, Appli
18	31	75.6	505	1 US-08-630-592-4	Sequence 4, Appli
19	31	75.6	505	1 US-08-714-991-4	Sequence 4, Appli
20	31	75.6	505	2 US-08-829-553-2	Sequence 2, Appli
21	31	75.6	505	2 US-08-922-267A-2	Sequence 2, Appli
22	31	75.6	505	2 US-08-936-707A-2	Sequence 2, Appli
23	31	75.6	505	2 US-08-936-706A-2	Sequence 2, Appli
24	31	75.6	505	3 US-09-248-203-2	Sequence 2, Appli
25	31	75.6	505	3 US-09-032-365A-4	Sequence 4, Appli
26	31	75.6	505	4 US-08-812-824-3	Sequence 3, Appli
27	31	75.6	505	4 US-09-406-071-2	Sequence 2, Appli

28 31 75.6 506 1 US-08-631-200-8 Sequence 8, Appli
29 31 75.6 506 2 US-08-829-553-8 Sequence 8, Appli
30 31 75.6 506 2 US-08-922-267A-8 Sequence 8, Appli
31 31 75.6 506 2 US-08-936-707A-8 Sequence 8, Appli
32 31 75.6 506 2 US-08-936-706A-8 Sequence 8, Appli
33 31 75.6 506 3 US-09-248-203-8 Sequence 8, Appli
34 31 75.6 506 3 US-09-032-365A-62 Sequence 62, Appli
35 31 75.6 506 4 US-08-812-824-4 Sequence 4, Appli
36 31 75.6 506 4 US-09-406-071-8 Sequence 8, Appli
37 31 75.6 512 3 US-09-032-365A-60 Sequence 60, Appli
38 31 75.6 518 3 US-09-032-365A-58 Sequence 58, Appli
39 31 75.6 561 1 US-08-714-991-27 Sequence 27, Appli
40 31 75.6 561 3 US-09-032-365A-10 Sequence 10, Appli
41 30 73.2 73 1 US-08-379-538-7 Sequence 7, Appli
42 30 73.2 452 2 US-08-731-079A-2 Sequence 2, Appli
43 30 73.2 547 4 US-09-178-252-11 Sequence 11, Appli
44 30 73.2 725 1 US-08-448-170-4 Sequence 4, Appli
45 30 73.2 725 4 US-08-961-803-7 Sequence 7, Appli

ALIGNMENTS

RESULT 1
US-07-620-410-3
; Sequence 3, Application US/07620410
; Patent No. 5217953
; GENERAL INFORMATION:
; APPLICANT: Gozes, Illana
; APPLICANT: Brennenman, Douglas E.
; APPLICANT: Fridkin, Mati
; APPLICANT: Moody, Terry
; TITLE OF INVENTION: A NOVEL VASOACTIVE INTESTINAL PEPTIDE
; TITLE OF INVENTION: ANTAGONIST
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
; STREET: 1615 L. Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/620,410
; FILING DATE: 19901130
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Scott, Watson T.
; REGISTRATION NUMBER: 26,581
; REFERENCE/DOCKET NUMBER: WTS/5683/82679
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 248453 CUSH
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
US-07-620-410-3

Query Match 97.6%; Score 40; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
|||||||

DB 2 MYPTLK 8

RESULT 2
US-08-922-267A-65
; Sequence 65, Application US/08922267A.
; Patent No. 5861239
; GENERAL INFORMATION:
; APPLICANT: Klevn, Patrick W.
; APPLICANT: Moore, Karen J.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; NUMBER OF SEQUENCES: 82
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/922,267A
; FILING DATE: 2-SEP-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/829,553
; FILING DATE: 28-MAR-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/631,200
; FILING DATE: 12-APR-1996
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-085
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-922-267A-65

Query Match 75.6%; Score 31; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 5
DB 2 MYPTY 6

RESULT 3
US-08-444-628-8
; Sequence 8, Application US/08444628
; Patent No. 6184021
; GENERAL INFORMATION:
; APPLICANT: Senior, Robert M.
; TITLE OF INVENTION: Catalytically-active Gelatinase Mutant
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Scott J. Meyer, Monsanto/Searle, A3SG
; STREET: 800 N. Lindbergh Blvd.
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63167
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/444,628
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/357,820
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyer, Scott J.
; REGISTRATION NUMBER: 25,275
; REFERENCE/DOCKET NUMBER: WU-2855
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)694-3117
; TELEFAX: (314)694-5435
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 173 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-444-628-8

Query Match 75.6%; Score 31; DB 4; Length 173;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
DB 138 MYPTY 142

RESULT 4
US-08-357-820-8
; Sequence 8, Application US/08357820
; Patent No. 6194189
; GENERAL INFORMATION:
; APPLICANT: Senior, Robert M.
; TITLE OF INVENTION: Catalytically-active Gelatinase Mutant
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scott J. Meyer, Monsanto/Searle, A3SG
; STREET: 800 N. Lindbergh Blvd.
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63167
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/357,820
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyer, Scott J.
; REGISTRATION NUMBER: 25,275
; REFERENCE/DOCKET NUMBER: WU-2855
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (314)694-3117
TELEFAX: (314)694-5435
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 173 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-357-820-8

Query Match 75.6%; Score 31; DB 4; Length 173;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
DB 138 MYPTY 142

RESULT 5

US-09-009-156-5
Sequence 5, Application US/09009156
Patent No. 6046031
GENERAL INFORMATION:
APPLICANT: Ni, Jian
APPLICANT: Ruben, Steven
APPLICANT: Brewer, Laurie
APPLICANT: Gentz, Reiner
TITLE OF INVENTION: No. 6046031el Metalloproteinases
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: US
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/009,156
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/034,205
FILING DATE: 21-JAN-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/049,607
FILING DATE: 13-JUN-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/054,541
FILING DATE: 01-AUG-1997
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PF376
TELEPHONE: 301-309-8504
TELEFAX: 301-309-8439
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 261 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-009-156-5

Query Match 75.6%; Score 31; DB 3; Length 261;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
DB 226 MYPTY 230

RESULT 6

US-09-009-156-6
Sequence 6, Application US/09009156
Patent No. 6046031
GENERAL INFORMATION:
APPLICANT: Ni, Jian
APPLICANT: Ruben, Steven
APPLICANT: Brewer, Laurie
APPLICANT: Gentz, Reiner
TITLE OF INVENTION: No. 6046031el Metalloproteinases
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: US
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/009,156
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/034,205
FILING DATE: 21-JAN-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/049,607
FILING DATE: 13-JUN-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/054,541
FILING DATE: 01-AUG-1997
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PF376
TELEPHONE: 301-309-8504
TELEFAX: 301-309-8439
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 264 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-009-156-6

Query Match 75.6%; Score 31; DB 3; Length 264;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
DB 232 MYPTY 236

RESULT 7

US-08-448-489-18
 ; Sequence 18, Application US/08448489
 ; Patent No. 6184022
 ; GENERAL INFORMATION:
 ; APPLICANT: SEIKI, Motoharu
 ; APPLICANT: SATO, Hiroshi
 ; APPLICANT: SHINAGAWA, Akira
 ; TITLE OF INVENTION: NOVEL METALLOPROTEINASE AND ENCODING DNA THEREFOR
 ; FILE REFERENCE: 55-290P
 ; CURRENT APPLICATION NUMBER: US/08/448,489
 ; CURRENT FILING DATE: 1995-06-07
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 18
 ; LENGTH: 267
 ; TYPE: PRT
 ; ORGANISM: Unknown
 ; FEATURE:
 ; OTHER INFORMATION: Description of Unknown Organism: Known Member of
 ; OTHER INFORMATION: Matrix Metalloproteinase Family
 ; 08-448-489-18

Query Match 75.6%; Score 31; DB 4; Length 267;
 Best Local Similarity 100.0%; Pred. No. 69;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 MYPTY 6
 Db 232 MYPTY 236

RESULT 8
 US-08-896-062-2
 ; Sequence 2, Application US/08896062
 ; Patent No. 6010893
 ; GENERAL INFORMATION:
 ; APPLICANT: KHIRA, Yasunori
 ; TITLE OF INVENTION: PROCESS FOR PRODUCING HUMAN MATRILYSIN
 ; TITLE OF INVENTION: BY MEANS OF RECOMBINANT DNA
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Browdy and Neimark
 ; STREET: 419 Seventh Street N.W., Ste. 300
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: United States of America
 ; ZIP: 20004
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/896,062
 ; FILING DATE: 17-JUL-1997
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/530,984
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Browdy, Roger L.
 ; REGISTRATION NUMBER: 25,618
 ; REFERENCE/DOCKET NUMBER: KHIRA-2
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 628-5197
 ; TELEFAX: (202) 737-3528
 ; TELEX: 248633
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 271 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear

US-08-896-062-2
 ; MOLECULE TYPE: protein

Query Match 75.6%; Score 31; DB 3; Length 271;
 Best Local Similarity 100.0%; Pred. No. 70;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 MYPTY 6
 Db 236 MYPTY 240

RESULT 9
 US-08-630-592-2
 ; Sequence 2, Application US/08630592
 ; Patent No. 5770432
 ; GENERAL INFORMATION:
 ; APPLICANT: Nishina, Patsy
 ; APPLICANT: No. 5770432entrauth, Konrad
 ; APPLICANT: Naggert, Jürgen
 ; APPLICANT: No. 5770432entr, Michael
 ; TITLE OF INVENTION: Obesity Associated Genes
 ; NUMBER OF SEQUENCES: 25
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
 ; STREET: 3400 Embarcadero Center, Suite 3400
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 941114187
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PCDOS/MSDOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/630,592
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Sherwood, Pamela J.
 ; REGISTRATION NUMBER: 36,677
 ; REFERENCE/DOCKET NUMBER: A59504/BIR/PJS
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 7811989
 ; TELEFAX: (415) 3983249
 ; TELEX: 910 277299
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 459 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-08-630-592-2

Query Match 75.6%; Score 31; DB 1; Length 459;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 MYPTY 6
 Db 233 MYPTY 237

RESULT 10
 US-08-714-991-2
 ; Sequence 2, Application US/08714991
 ; Patent No. 5776762
 ; GENERAL INFORMATION:
 ; APPLICANT: NORTH, Michael

APPLICANT: NISHINA, Patsy
APPLICANT: No. 5776762en-Trauth, Konrad
APPLICANT: NAGGERT, Juergen
TITLE OF INVENTION: OBESITY ASSOCIATED GENES
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/714,991
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: SHERWOOD, Pamela J.
REGISTRATION NUMBER: 36,677
REFERENCE/DOCKET NUMBER: A-59504-1/PJS
TELEPHONE: 415-494-8700
TELEFAX: 415-494-8771
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 459 amino acids
TYPE: amino acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-714-991-2

Query Match 75.6%; Score 31; DB 1; Length 459;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
|||||
Db 233 MYPTY 237

RESULT 11
US-09-032-365A-2
Sequence 2, Application US/09032365A
Patent No. 6114502
GENERAL INFORMATION:
APPLICANT: No. 6114502th, Michael
APPLICANT: Nishina, Patsy
APPLICANT: Naggart, Juergen
APPLICANT: No. 6114502en-Trauth, Konrad
TITLE OF INVENTION: GENE FAMILY ASSOCIATED WITH
TITLE OF INVENTION: NEUROSENSORY DEFECTS
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bozicevic & Reed, LLP
STREET: 285 Hamilton Avenue, Suite 200
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/032,365A
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Sherwood, Pamela J.
REGISTRATION NUMBER: 36,677
REFERENCE/DOCKET NUMBER: SEQ-2CIP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-327-3400
TELEFAX: 650 327-3231
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 459 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-032-365A-2

Query Match 75.6%; Score 31; DB 3; Length 459;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
|||||
Db 233 MYPTY 237

RESULT 12
US-08-630-592-7
Sequence 7, Application US/08630592
Patent No. 5770432
GENERAL INFORMATION:
APPLICANT: Nishina, Patsy
APPLICANT: No. 5770432enTrauth, Konrad
APPLICANT: Naggert, Juergen
APPLICANT: No. 5770432th, Michael
TITLE OF INVENTION: Obesity Associated Genes
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
STREET: 3400 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 941114187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PCDOS/MSDOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/630,592
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sherwood, Pamela J.
REGISTRATION NUMBER: 36,677
REFERENCE/DOCKET NUMBER: A59504/BIR/PJS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 7811989
TELEFAX: (415) 3983249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 460 amino acids
TYPE: amino acid
STRANDEDNESS: double

TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-630-592-7

Query Match 75.6%; Score 31; DB 1; Length 460;
Best Local Similarity 100.0%; Pred. NO. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
Db 234 MYPTY 238

RESULT 13
US-08-714-991-7
Sequence 7, Application US/08714991
Patent No. 5776762
GENERAL INFORMATION:
APPLICANT: NISHINA, Michael
APPLICANT: NISHINA, Patsy
APPLICANT: No. 5776762en-Trauth, Konrad
TITLE OF INVENTION: OBESITY ASSOCIATED GENES
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, ROHRBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/714, 991
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: SHERWOOD, Pamela J.
REGISTRATION NUMBER: 36,677
REFERENCE/DOCKET NUMBER: A-59504-1/PJS
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-494-8770
TELEFAX: 415-494-8771
TELEX: 910 277299

INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 460 amino acids
TYPE: amino acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-714-991-7

Query Match 75.6%; Score 31; DB 1; Length 460;
Best Local Similarity 100.0%; Pred. NO. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
Db 234 MYPTY 238

RESULT 14
US-09-032-365A-8
Sequence 8, Application US/09032365A
Patent No. 6114502
GENERAL INFORMATION:

APPLICANT: No. 6114502th, Michael
APPLICANT: Nishina, Patsy
APPLICANT: Naggart, Juergen
APPLICANT: No. 6114502en-Trauth, Konrad
TITLE OF INVENTION: GENE FAMILY ASSOCIATED WITH
TITLE OF INVENTION: NEUROSENSORY DEFECTS
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bozicevic & Reed, LLP
STREET: 285 Hamilton Avenue, Suite 200
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PC-DOS for Windows Version 2.0
CURRENT APPLICATION DATA: US/09/032, 365A
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Sherwood, Pamela J.
REGISTRATION NUMBER: 36,677
REFERENCE/DOCKET NUMBER: SEQ-2CIP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-327-3400
TELEFAX: 650-327-3231
TELEX:

INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 460 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-032-365A-8

Query Match 75.6%; Score 31; DB 3; Length 460;
Best Local Similarity 100.0%; Pred. NO. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
Db 234 MYPTY 238

RESULT 15
US-08-068-392-3
Sequence 3, Application US/08068392
Patent No. 6150152
GENERAL INFORMATION:
APPLICANT: Shuprio, Steven M.
TITLE OF INVENTION: Human Macrophage Metalloproteinase
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scott J. Meyer, Monsanto Co., A3SM
STREET: 800 N. Lindbergh Blvd.
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63167
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/068,392
FILING DATE: 19930528
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Meyer, Scott J.
REGISTRATION NUMBER: 25275
REFERENCE/DOCKET NUMBER: 07-24(12406)A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314)694-3117
TELEFAX: (314)694-5435
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 462 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-068-392-3

```

Query Match	75.5%;	Score 31;	DB 4;	Length 462;
Best Local Similarity	100.0%;	Pred. NO. 1.2e+02;		
Matches 5;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	2 MYTY 6
D _b	229 MYPTY 233

Search completed: September 4, 2001, 15:51:13
Job time: 196 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein.- protein search, using sw model

Run on: September 4, 2001, 15:52:09 ; Search time 49.08 Seconds
(without alignments)
12.416 Million cell updates/sec

Title: US-09-630-345-5

Perfect score: 41

Sequence: 1 XMYPTYLK 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

al number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_68:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	37	90.2	384	2 I51267	collagenase (EC 3.4.24.-) - bullfrog
2	35	85.4	393	2 B86644	transporter ybfb [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
3	34	82.9	436	2 C69764	4-aminobutyrate am
4	33	80.5	218	2 E86750	hypothetical prote
5	33	80.5	269	2 D69011	hypothetical prote
6	32	78.0	340	1 MMBE5	cell fusion protei
7	32	78.0	349	2 JU0458	peroxidase (EC 1.1
8	32	78.0	593	2 JC4884	organic cation tra
9	32	78.0	1263	2 F44466	DNA-directed RNA p
10	31	75.6	247	2 B85515	hypothetical prote
11	31	75.6	262	2 T29698	hypothetical prote
12	31	75.6	267	1 KCHUM	matrilysin (EC 3.4
13	31	75.6	267	2 A57490	matrilysin (EC 3.4
14	31	75.6	301	2 S57923	SEC14 protein - ye
15	31	75.6	378	2 F86307	hypothetical prote
16	31	75.6	385	2 E86359	Similar to seed ma
17	31	75.6	462	2 A42401	macrophage elastas
18	31	75.6	478	2 A86677	phospho-beta-gluc
19	31	75.6	483	2 JCS743	matrix metalloprot
20	31	75.6	495	1 P1WLB	L1 protein - bovin
21	31	75.6	501	1 P1WLB2	L1 protein - bovin
22	31	75.6	501	1 P1WLEP	L1 protein - Europ
23	31	75.6	502	2 T13708	hypothetical prote
24	31	75.6	503	2 T29175	hypothetical prote
25	31	75.6	505	2 S68518	tub protein, brain
26	31	75.6	511	2 T19496	hypothetical prote
27	31	75.6	513	1 P1WLDP	L1 protein - deer
28	31	75.6	516	2 T15633	hypothetical prote
29	31	75.6	565	2 H69363	hypothetical prote

30	75.6	662	2 S42826	probable ATPase -
31	75.6	2185	1 GNNYSV	genome polyprotein
32	75.6	2185	1 GNNYSV	genome polyprotein
33	73.2	141	2 T29507	hypothetical prote
34	73.2	152	2 G83476	hypothetical prote
35	73.2	155	2 T17986	hypothetical prote
36	73.2	192	2 T11930	NADH dehydrogenase
37	73.2	239	2 B64757	glycolate oxidase
38	73.2	239	2 D85522	probable dehydrog
39	73.2	248	2 S57910	probable response
40	73.2	262	2 C81384	shikimate 5-dehydr
41	73.2	287	2 D83414	probable transcrip
42	73.2	343	1 MMBEAS	cell fusion protei
43	73.2	368	1 HLBECM	membrane glycoprot
44	73.2	410	2 T20397	hypothetical prote
45	73.2	491	2 JE0396	phospho-beta-galac

ALIGNMENTS

RESULT 1
I51267
collagenase (EC 3.4.24.-) - bullfrog
C:Species: Rana catesbeiana (bullfrog)
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 22-Jun-1999
C:Accession: I51267
R:Oofusa, K.; Yomori, S.; Yoshizato, K.
Int. J. Dev. Biol. 38, 345-350, 1994
A:Title: Regionally and hormonally regulated expression of genes of collagen and coll
A:Reference number: I51267; MUID:95071832
A:Accession: I51267
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-384 <OOF>
A:Cross-references: GB:S75623; NID:9913070; PIDN:AAB32661.1; PID:9913071
A:Superfamily: interstitial collagenase; hemopexin repeat homology; matrix metallopro
F:Keywords: hydrolase; metalloproteinase; zinc; zymogen
F:536-231/Domain: matrix metalloproteinase homology <MMP>
F:236-381/Domain: hemopexin repeat homology <PN>
F:81,189,193,199/Binding site: zinc, catalytic (Cys, His, His, His) (inhibited) #stat
F:189,193,199/Binding site: zinc, catalytic (His) (active) #status predicted
F:190/Active site: Glu #status predicted

Query Match 90.2%; Score 37; DB 2; Length 384;
Best Local Similarity 85.7%; Pred. No. 5;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
|||||
DB 207 MYPTYLK 213

RESULT 2
B86644
transporter ybfb [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 23-Mar-2001
C:Accession: B86644
R:Polotin, A.; Winkler, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Eh
Genome Res. in press, 2001
A:Title: The complete genome sequence of the lactic acid bacterium.
A:Reference number: A86625
A:Accession: B86644
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-393 <STO>
A:Cross-references: GB:AE005176; NID:gi12723004; PIDN:AAK04252.1; GSPDB:GN00146
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: ybfb

Query Match 85.4%; Score 35; DB 2; Length 393;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYL 7
 |||||
 DB 278 MYPTYL 283

RESULT 3
 C69764
 4-aminobutyrate aminotransferase homolog yncG - Bacillus subtilis
 C:Species: Bacillus subtilis
 C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
 C:Accession: C69764
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berteaux, C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chao, A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.; Fure 390, 249-256, 1997
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallen, J.; Harwood, C.R.; Hensaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F.; Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, Y. M.; Ogawa, K.; Ogiwara, K.; Oudega, B.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel, R.; Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon, A.; Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron, akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
 A:Reference number: A69580; MUID:98044033
 A:Accession: C69764
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-436 <KUN>
 A:Cross-references: GB:Z93106; GB:AL009126; NID:g2632653; PIDN:CAB12198.1; PID:g2632691
 A:Experimental source: strain 168
 C:Genetics:
 A:Gene: yncG
 C:Superfamily: ornithine-oxo-acid aminotransferase

Query Match 82.9%; Score 34; DB 2; Length 436;
 Best Local Similarity 71.4%; Pred. No. 23;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYL 8
 |||||
 DB 88 MYPTYL 94

RESULT 4
 E86750
 hypothetical protein ykdB [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
 C:Species: Lactococcus lactis subsp. lactis
 C>Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 23-Mar-2001
 C:Accession: E86750
 R:Bohlin, A.; Winkler, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissensbach, J.; Ehrlich, Genome Res. in press, 2001
 A:Title: the complete genome sequence of the lactic acid bacterium.
 A:Reference number: A86625
 A:Accession: E86750
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-218 <STO>
 A:Cross-references: GB:AE005176; NID:q12723952; PIDN:AAK05103.1; GSPDB:GN00146
 A:Experimental source: strain IL1403
 C:Genetics:
 A:Gene: ykdB

Query Match 80.5%; Score 33; DB 2; Length 218;

QY 2 MYPTYL 8
 |||||
 DB 244 LYPTYIR 250

RESULT 5
 D69011
 hypothetical protein MTH1087 - Methanobacterium thermoautotrophicum (strain Delta H)
 C:Species: Methanobacterium thermoautotrophicum
 C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
 C:Accession: D69011
 R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, K.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
 J. Bacteriol. 179, 7135-7155, 1997
 A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: fu
 A:Reference number: A69000; MUID:98037514
 A:Accession: D69011
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-269 <MTH>
 A:Cross-references: GB:AE000879; GB:AE000666; NID:g2622175; PIDN:AAB85576.1; PID:g262
 A:Experimental source: strain Delta H
 C:Genetics:
 A:Gene: MTH1087

Query Match 80.5%; Score 33; DB 2; Length 269;
 Best Local Similarity 83.3%; Pred. No. 23;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 YPTYLK 8
 |||||
 DB 175 YPTYIK 180

RESULT 6
 MMB5
 cell fusion protein precursor - human herpesvirus 3
 C:Species: human herpesvirus 3, varicella-zoster virus
 C>Date: 30-Sep-1988 #sequence_revision 30-Sep-1988 #text_change 16-Jul-1999
 C:Accession: E27212
 R:Davidson, A.J.; Scott, J.E.
 J. Gen. Virol. 67, 1759-1816, 1986
 A:Title: The complete DNA sequence of varicella-zoster virus.
 A:Reference number: A27345; MUID:86306657
 A:Accession: E27212
 A:Molecule type: DNA
 A:Residues: 1-340 <DAV>
 A:Cross-references: EMBL:X04370; NID:g59989; PIDN:CAA27888.1; PID:g59994
 C:Genetics:
 A:Gene: 5
 C:Superfamily: herpesvirus cell fusion protein
 C:Keywords: membrane fusion; transmembrane protein
 F:1-31/Domain: signal sequence #status predicted <SIG>
 F:32-340/Product: cell fusion protein #status predicted <CFF>
 F:115-137/Domain: transmembrane #status predicted <TM1>
 F:220-238/Domain: transmembrane #status predicted <TM2>
 F:251-269/Domain: transmembrane #status predicted <TM3>
 F:307-322/Domain: transmembrane #status predicted <TM5>

Query Match 78.0%; Score 32; DB 1; Length 340;
 Best Local Similarity 57.1%; Pred. No. 47;
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYL 8
 |||||
 DB 244 LYPTYIR 250

RESULT 7

JU0458
 peroxidase (EC 1.11.1.7) E - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 04-Mar-2000
 C:Accession: JU0458
 R:Incepruk, C.; Higashimura, N.; Yamamoto, K.; Okada, N.; Shimmyo, A.; Takano, M.
 Gene 98, 237-241, 1991
 A:Title: Nucleotide sequences of two genomic DNAs encoding peroxidase of Arabidopsis thaliana
 A:Reference number: JU0457; MUID:91200671
 A:Accession: JU0458
 A:Molecule type: DNA
 A:Residues: 1-349 <INT>
 A:Cross-references: GB:M58381; NID:gl66806; PIDN:AAA32842.1; PID:gl66807
 C:Genetics:
 A:Gene: prxEa
 A:Introns: 76/3; 140/3; 197/1
 C:Superfamily: peroxidase
 Keywords: chromoprotein; heme; iron; metalloprotein; oxidoreductase
 F:10-120/Disulfide bonds: #status predicted
 F:57/Active site: Arg #status predicted
 F:71,199/Binding site: heme iron (His) (axial ligands) #status predicted
 F:73-78/Disulfide bonds: #status predicted
 F:126-329/Disulfide bonds: #status predicted
 F:206-238/Disulfide bonds: #status predicted

Query Match 78.0%; Score 32; DB 2; Length 349;
 Best Local Similarity 83.3%; Pred. No. 48;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYL 7

Db 226 LYPTYL 231
 :|||||

RESULT 8

JC4884
 organic cation transporter protein 2 - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 10-Sep-1996 #sequence_revision 18-Oct-1996 #text_change 05-Nov-1999
 C:Accession: JC4884
 R:Okuda, M.; Saito, H.; Urakami, Y.; Takano, M.; Inui, K.
 Biochem. Biophys. Res. Commun. 224, 500-507, 1996
 A:Title: cDNA cloning and functional expression of a novel rat kidney organic cation transporter
 A:Reference number: JC4884; MUID:96295517
 A:Accession: JC4884
 A:Molecule type: mRNA
 A:Residues: 1-593 <OKU>
 A:Cross-references: DBJ:DR3044; NID:gl502282; PIDN:RAA11754.1; PID:dl012421; PID:gl5022
 A:Experimental source: kidney
 C:Comment: This protein is responsible for the transport of cationic drugs in kidney.

Query Match 78.0%; Score 32; DB 2; Length 593;
 Best Local Similarity 57.1%; Pred. No. 82;
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYLK 8

Db 457 LYPTIIR 463
 :|||||

RESULT 9

F44466
 DNA-directed RNA polymerase (EC 2.7.7.6) beta chain - Thermotoga maritima (strain MSB8)
 C:Species: Thermotoga maritima
 C:Date: 10-Jun-1993 #sequence_revision 26-Apr-1996 #text_change 21-Jul-2000
 C:Accession: S41466; S41462; F44466; B72373; S19903
 R:Palm, P.; Schleper, C.; Arnold-Ammer, I.; Holz, I.; Meier, T.; Lottspeich, F.; Zillig, W.
 Nucleic Acids Res. 21, 4904-4908, 1993
 A:Title: The DNA-dependent RNA-polymerase of Thermotoga maritima; characterisation of the

A:Reference number: S41462; MUID:94232816
 A:Accession: S41466
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-1263 <PAL>
 A:Cross-references: EMBL:X72695; NID:g425255; PIDN:CAA51246.1; PID:g425258
 A:Experimental source: DSM 3109
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1993
 A:Accession: S41462
 A:Molecule type: protein
 A:Residues: 1-5 'X', 7-23 <PAW>
 A:Experimental source: DSM 3109
 R:Liao, D.; Dennis, P.P.
 J. Biol. Chem. 267, 22787-22797, 1992
 A:Title: The organization and expression of essential transcription translation components
 A:Reference number: A44466; MUID:93054590
 A:Accession: F44466
 A:Molecule type: DNA
 A:Residues: 1-404 <LIA>
 A:Cross-references: EMBL:Z11839; NID:g407020; PIDN:CAA77863.1; PID:g48188
 A:Note: sequence extracted from NCBI backbone (NCBIP:118059)
 R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hic
 Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, C.M.
 Nature 399, 323-329, 1999
 A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome
 A:Reference number: A72200; MUID:99287316
 A:Accession: B72373
 A:Molecule type: DNA
 A:Residues: 1-1263 <ARN>
 A:Cross-references: GB:AE001724; GB:AE000512; NID:g4980966; PIDN:AAD35543.1; PID:g498
 A:Experimental source: strain MSB8
 C:Genetics:
 A:Gene: rpoB
 C:Superfamily: DNA-directed RNA polymerase beta chain
 C:Keywords: nucleotidyltransferase; transcription
 F:1-1263/Product: DNA-directed RNA polymerase beta chain #status predicted <MAT>

Query Match 78.0%; Score 32; DB 2; Length 1263;
 Best Local Similarity 83.3%; Pred. No. 1.8e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYL 7

Db 214 LYPTYL 219
 :|||||

RESULT 10

B85515
 hypothetical protein 20321 [Imported] - Escherichia coli (strain O157:H7)
 C:Species: Escherichia coli
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 31-Mar-2001
 R:Ferna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May
 Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimaienta, E.; Potamousis, K.; Apoda
 Nature 409, 529-533, 2001
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A:Reference number: B85480; MUID:21074935; PMID:11206551
 A:Accession: B85515
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-247 <STO>
 A:Cross-references: GB:AE005174; NID:gl2513034; PIDN:AAG54582.1; GSPDB:GN00145; UWGP:
 A:Experimental source: strain O157:H7, substrain EDL933
 C:Genetics:
 A:Gene: 20321
 C:Superfamily: hypothetical protein f242a

Query Match 75.6%; Score 31; DB 2; Length 247;
 Best Local Similarity 83.3%; Pred. No. 54;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 YPTYLK 8
|:|:|:|
Db 76 YPSYLK 81

RESULT 11

T29698
hypothetical protein F31A3.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T29698

Murray, J.; Le, T.T.

Submitted to the EMBL Data Library, May 1996

A:Description: The sequence of C. elegans cosmid F31A3.

A:Reference number: Z20667

A:Accession: T29698

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-262 <NR>

A:Cross-references: EMBL:U58742; PIDN:AAB36855.1; GSPDB:GN00028; CESP:F31A3.2

C:Experimental source: strain Bristol N2; clone F31A3

A:Gene: CESP:F31A3.2

A:Map position: X

A:Introns: 23/3; 80/2; 155/2

Query Match 75.6%; Score 31; DB 2; Length 262;

Best Local Similarity 71.4%; Pred. No. 58;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 MYPTYLK 8

|:|:|:|

Db 188 MYPVILE 194

RESULT 12

KCHUM

matrilysin (EC 3.4.24.23) precursor - human

N:Alternate names: matrix metalloproteinase 7 (MMP7); probable metalloproteinase

N:Contains: promatrilysin

C:Species: Homo sapiens (man)

C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 18-Jun-1999

C:Accession: B28816; A60539; S24324

R:Muller, D.; Quantin, B.; Genest, M.C.; Millon-Collard, R.; Abecassis, J.; Breathnach,

Biochem. J. 253, 187-192, 1988

A:Title: The collagenase gene family in humans consists of at least four members.

A:Reference number: A90339; MUID:88339885

A:Accession: B28816

A:Molecule type: mRNA

A:Residues: 1-267 <MUL>

A:Cross-references: EMBL:X07819; NID:g35798; PIDN:CAA30678.1; PID:g35799

R:Miyazaki, K.; Hattori, Y.; Umenishi, F.; Yasumitsu, H.; Umeda, M.

Cancer Res. 50, 7758-7764, 1990

A:Title: Purification and characterization of extracellular matrix-degrading metallopro-

A:Reference number: A60539; MUID:91070531

A:Accession: A60539

A:Molecule type: protein

A:Residues: 18-35, 'X', 37-42 <MIY>

R:Marti, H.P.; McNeil, L.; Thomas, G.; Davies, M.; Lovett, D.H.

Biochem. J. 285, 899-905, 1992

A:Title: Molecular characterization of a low-molecular-mass matrix metalloproteinase sec

A:Reference number: S24324; MUID:92359961

A:Accession: S24324

A:Molecule type: DNA

A:Status: preliminary

A:Residues: 1-267 <NR>

A:Cross-references: EMBL:Z11887; NID:g35802; PIDN:CAA77942.1; PID:g35803

C:Comment: This enzyme is similar in its activity to stromelysin and degrades various ex

s of types II, IV, IX, X, and XI.

C:Comment: Matrilysin hydrolyzes peptide bonds in plasminogen to yield a fragment with a

C:Genetics:

A:Gene: GDB:MMP7; MPSL1
A:Cross-references: GDB:125751; OMIM:178990
A:Map position: 11q21-11q22
C:Superfamily: matrilysin; matrix metalloproteinase homology
C:Keywords: calcium; extracellular matrix; fibroblast; hydrolase; metalloproteinase;
F:1-17/Domain: signal sequence #status predicted <SIG>
F:18-267/Product: promatrilysin #status predicted <PRO>
F:18-94/Domain: activation peptide #status predicted <ACT>
F:55-259/Domain: matrix metalloproteinase homology <MMP>
F:85-92/Region: autoinhibitory
F:95-267/Product: matrilysin #status predicted <MAT>
F:87,214,218,224/Binding site: zinc, catalytic (Cys, His, His) (inhibited) #stat
F:214,218,224/Binding site: zinc, catalytic (His) (active) #status predicted
F:215/Active site: Glu #status predicted

Query Match 75.6%; Score 31; DB 1; Length 267;

Best Local Similarity 100.0%; Pred. No. 59;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6

|:|:|:|

Db 232 MYPTY 236

RESULT 13

A57490

matrilysin (EC 3.4.24.23) precursor - rat

N:Alternate names: matrix metalloproteinase 7 (MMP7)

C:Species: Rattus norvegicus (Norway rat)

C:Date: 08-Dec-1995 #sequence_revision 08-Dec-1995 #text_change 22-Jun-1999

C:Accession: A57490

R:Abramson, S.R.; Conner, G.E.; Nagase, H.; Neuhaus, I.; Woessner Jr., J.F.

J. Biol. Chem. 270, 16016-16022, 1995

A:Title: Characterization of rat uterine matrilysin and its cDNA. Relationship to hum

A:Reference number: A57490; MUID:95332299

A:Accession: A57490

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-267 <ABR>

A:Cross-references: GB:I24374; NID:9402492; PIDN:AAA99432.1; PID:g402493

C:Superfamily: matrilysin; matrix metalloproteinase homology

C:Keywords: hydrolase; metalloproteinase; zinc; zymogen

F:1-20/Domain: signal sequence #status predicted <SIG>

F:21-267/Product: matrilysin #status predicted <MAT>

F:58-262/Domain: matrix metalloproteinase homology <MMP>

F:90,217,221,227/Binding site: zinc, catalytic (Cys, His, His) (inhibited) #stat

F:217,221,227/Binding site: zinc, catalytic (His) (active) #status predicted

F:218/Active site: Glu #status predicted

Query Match 75.6%; Score 31; DB 2; Length 267;

Best Local Similarity 100.0%; Pred. No. 59;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6

|:|:|:|

Db 235 MYPTY 239

RESULT 14

S57923

SEC14

protein - yeast (Candida albicans)

C:Species: Candida albicans

C:Date: 13-Jun-1999 #sequence_revision 01-Mar-1996 #text_change 26-Feb-1998

C:Accession: S57923; S72193

R:Montoliu, L.; Sanz, M.; Pla, J.; Gil, C.; Nombela, C.

submitted to the EMBL Data Library, September 1994

A:Description: Characterisation of the Candida albicans SEC14 homolog gene.

A:Reference number: S57923

A:Accession: S57923

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-301 <MON>
 A:Cross-references: EMBL:X81937
 R:Montecoliva, L.; Sanchez, M.; Pla, J.; Gil, C.; Nombela, C.
 Feast 12, 1097-1105, 1996
 A:Title: Cloning of Candida albicans SEC14 gene homologue coding for a putative essential protein
 A:Reference number: S72193; MUID:97051600
 A:Accession: S72193
 A:Molecule type: DNA
 A:Residues: 1-301 <MON>
 A:Cross-references: EMBL:X81937
 A:Note: the authors translated the codon CTG for residue 180 as Ser
 C:Genetics:
 A:Gene: SEC14
 C:Superfamily: cellular retinaldehyde-binding protein; cellular retinaldehyde-binding protein
 F:59-266/Domain: cellular retinaldehyde-binding protein homology <CRB>

Query Match 75.6%; Score 31; DB 2; Length 301;
 Best Local Similarity 100.0%; Pred. No. 66;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2 MYPTY 6
 |||||
 Db 106 MYPTY 110

RESULT 15
 F86307
 hypothetical protein AAD50017.1 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: F86307
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719
 A:Accession: F86307
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-378 <STO>
 A:Cross-references: GB:AE005172; NID:g5734752; PIDN:AAD50017.1; GSPDB:GN00141
 C:Genetics:
 Gap position: 1

Query Match 75.6%; Score 31; DB 2; Length 378;
 Best Local Similarity 100.0%; Pred. No. 84;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 MYPTY 6
 |||||
 Db 118 MYPTY 122

Search completed: September 4, 2001, 15:52:09
 Job time: 232 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2001, 15:54:09 ; Search time 27.19 Seconds
(without alignments)
10.079 Million cell updates/sec

Title: US-09-630-345-5
Perfect score: 41
Sequence: 1 XMPYLYK 8

Scoring table: BLOSOM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37	90.2	384	MM01_RANCA	Q11133 rana catesb
2	34	82.9	436	GABT_BACSU	P94427 bacillus su
3	32	78.0	340	CELF_VZVD	P09261 varicella-z
4	32	78.0	349	PERE_ARATH	P24102 arabidopsis
5	32	78.0	594	NU5M_HIPAM	Q92371 hippopotamu
6	32	78.0	1263	RPOB_THEMA	P29398 thermotoga
7	31	75.6	262	MM07_FELCA	P55032 felis silve
8	31	75.6	264	MM07_MOUSE	Q10738 mus musculu
9	31	75.6	267	MM07_HUMAN	P09237 homo sapien
10	31	75.6	267	MM07_RAT	P50280 rattus norv
11	31	75.6	301	SC14_CANAL	P46250 candida alb
12	31	75.6	462	MM12_MOUSE	P34960 mus musculu
13	31	75.6	465	MM12_RAT	P63341 rattus norv
14	31	75.6	467	MM18_XENLA	Q13065 xenopus lae
15	31	75.6	482	MM20_MOUSE	P57748 mus musculu
16	31	75.6	483	MM20_HUMAN	O60882 homo sapien
17	31	75.6	483	MM20_PIG	P79287 sus scrofa
18	31	75.6	495	VLL1_BPV1	P03103 bovine papi
19	31	75.6	497	VLL1_BPV2	P06458 bovine papi
20	31	75.6	501	VLL1_PAPVE	P11326 european el
21	31	75.6	503	YPTJ_CAEEL	Q23459 caenorhabdi
22	31	75.6	505	TUB_MOUSE	P50586 mus musculu
23	31	75.6	505	TUB_RAT	O88808 rattus norv
24	31	75.6	506	TUB_HUMAN	P50607 homo sapien
25	31	75.6	511	YELJ_CAEEL	P90756 caenorhabdi
26	31	75.6	513	VLL1_PAPVD	P03104 deer papill
27	31	75.6	662	YMEI_SCHMA	P46508 schistosoma
28	31	75.6	2185	POLG_SVDVH	P16604 s genome po
29	31	75.6	2185	POLG_SVDVU	P13900 s genome po
30	30	73.2	91	Y13F_BPT4	P39499 bacterioph
31	30	73.2	192	NUGM_PROWI	O37622 prototheca
32	30	73.2	239	YKGE_ECOLI	P77252 escherichia
33	30	73.2	308	Y04O_BPT4	P39254 bacterioph

34 30 73.2 343 1 CELF_HSVB P28933 equine herp
35 30 73.2 368 1 VGH3_HCMVA P08560 human cytom
36 30 73.2 425 1 BIOA_SERMA P36568 serratia ma
37 30 73.2 580 1 P69_MYCHR P15362 mycoplasma
38 30 73.2 634 1 VC09_VACCC P21042 vaccinia vi
39 30 73.2 634 1 VC09_VACCV P17372 vaccinia vi
40 30 73.2 1056 1 DPOL_ADE02 P03261 human adeno
41 30 73.2 1056 1 DPOL_ADE05 P04495 human adeno
42 30 73.2 1138 1 C7AA_BACTU Q03749 bacillus th
43 30 73.2 1138 1 C7AB_BACUA Q45707 bacillus th
44 30 73.2 1138 1 C7AB_BACUK Q45708 bacillus th
45 30 73.2 1155 1 C1AB_BACTK P06578 bacillus th

ALIGNMENTS

RESULT 1
MM01_RANCA STANDARD; PRT; 384 AA.
AC Q11133;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE INTERSTITIAL COLLAGENASE PRECURSOR (EC 3.4.24.7) (MATRIX
DE METALLOPROTEINASE-1) (MMP-1) (TCL).
OS Rana catesbeiana (Bull frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Ranidae; Rana.
OX NCBI_TaxID=8400;
RN [.]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RX MEDLINE=95071832; PubMed=7981043;
RA Oofusa K., Yomori S., Yoshizato K.;
RT "Regionally and hormonally regulated expression of genes of collagen
and collagenase in the anuran larval skin.";
RL Int. J. Dev. Biol. 38:345-350(1994).
CC -|- FUNCTION: CLEAVES COLLAGENS OF TYPES I, II, AND III AT ONE SITE IN
THE HELICAL DOMAIN. ALSO CLEAVES COLLAGENS OF TYPES VII AND X.
CC -|- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY.
CC -|- ENZYME REGULATION: CAN BE ACTIVATED WITHOUT REMOVAL OF THE
ACTIVATION PEPTIDE (BY SIMILARITY).
CC -|- SIMILARITY: CONTAINS 1 HEMOXEPIN-LIKE DOMAIN.
CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC
METALLOPROTEASE) ALSO KNOWN AS MATRIXIN SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

CC EMBL; S75623; AAB32661.1; -
CC HSP; P21692; IPBL.
DR MEROPS; M10.001;
DR InterPro; IPR000130;
DR InterPro; IPR000585;
DR InterPro; IPR001818;
DR Pfam; PF00413; Peptidase_M10; 1.
DR Pfam; PF00045; hemoxepin; 2.
DR PROSITE; PS00024; ZINC_PROTEASE; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
DR PROSITE; PS00546; CYSTEINE_SWITCH; FALSE_NEG.
KW Hydroxylase; Metalloprotease; zinc; Zymogen; Calcium;
Collagen degradation; Extracellular matrix; Signal.
FT SIGNAL 1 25 POTENTIAL.
FT PROPEP 26 88 ACTIVATION PEPTIDE (POTENTIAL).
FT CHAIN 89 384 INTERSTITIAL COLLAGENASE.
FT DOMAIN 239 384 HEMOXEPIN-LIKE.
FT SITE 81 81 CYSTEINE SWITCH (POTENTIAL).

```

FT METAL 189 189 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 190 190 BY SIMILARITY.
FT METAL 193 193 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 199 199 ZINC (CATALYTIC) (BY SIMILARITY).
FT DISULFID 242 381 PROBABLE.
SQ SEQUENCE 384 AA; 43582 MW; A5B5E2FB332239DF CRC64;

Query Match 90.2%; Score 37; DB 1; Length 384;
Best Local Similarity 85.7%; Pred. No. 2.9;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
DB 207 MYPTYLK 213

RESULT 2
GABT_BACSU STANDARD; PRT; 436 AA.
P94427;
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PROBABLE 4-AMINO BUTYRATE AMINOTRANSFERASE (EC 2.6.1.19) (GAMMA-AMINO-
DE N-BUTYRATE TRANSAMINASE) (GABA TRANSAMINASE) (GLUTAMATE-SUCCINIC
DE SEMIALDEHYDE TRANSAMINASE) (GABA AMINOTRANSFERASE) (GABA-AT).
GN GABT.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA Yamane K., Kumano M., Kurita K.;
RX MEDLINE=97124189; PubMed=8969502;
RT "The 25 degrees-36 degrees region of the Bacillus subtilis
RT chromosome: determination of the sequence of a 146 kb segment and
RT identification of 113 genes."
RL Microbiology 142:3047-3056(1996).
CC -!- CATALYTIC ACTIVITY: 4-AMINO BUTANOATE + 2-OXOGLUTARATE = SUCCINATE
CC -!- SEMIALDEHYDE + L-GLUTAMATE.
CC -!- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -!- PATHWAY: 4-AMINO BUTYRATE (GABA) DEGRADATION PATHWAY.
CC -!- SIMILARITY: BELONGS TO CLASS-III OF PYRIDOXAL-PHOSPHATE-DEPENDENT
CC AMINOTRANSFERASES.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

DR EMBL; D50453; BAA09021.1; -
DR EMBL; 299106; CAB12198.1; -
DR Subtilist; BG12043; gabt.
DR InterPro; IPR000954; -
DR Pfam; PF00202; aminotran_3; 1.
DR PROSITE; PS00600; AA_TRANSFPR_GIASS_3; 1.
KW Transferase; Aminotransferase; Pyridoxal phosphate.
FT BINDING 281 281 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ SEQUENCE 436 AA; 47249 MW; D0961F6D4189A8F3 CRC64;

Query Match 82.9%; Score 34; DB 1; Length 436;
Best Local Similarity 71.4%; Pred. No. 13;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
DB 207 MYPTYLK 213

```

```

DB 88 MYPTVIE 94

RESULT 3
CELLF_VZVD STANDARD; PRT; 340 AA.
AC P09261;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE CELL FUSION PROTEIN PRECURSOR.
GN 5.
OS Varicella-zoster virus (strain Dumas) (VZV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10338;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86306657; PubMed=3018124;
RA Davison A.J., Scott J.E.;
RT "The complete DNA sequence of varicella-zoster virus.";
RL J. Gen. Virol. 67:1759-1816(1986).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; X04370; CAA27888.1; -
DR PIR; E27212; MMBE5
DR InterPro; IPR002567; -
DR Pfam; PF01621; Fusion_gly_k; 1.
KW Fusion protein; Transmembrane; Signal.
FT SIGNAL 1 340
FT CHAIN 2 340 CELL FUSION PROTEIN.
FT CARBOHYD 55 55 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 65 65 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 340 AA; 38576 MW; 0387FE0UEC39C946 CRC64;

Query Match 78.0%; Score 32; DB 1; Length 340;
Best Local Similarity 57.1%; Pred. No. 25;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
DB 244 LYPTVIR 250

RESULT 4
PERE_ARATH STANDARD; PRT; 349 AA.
AC P24102;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE BASIC PEROXIDASE E PRECURSOR (EC 1.11.1.7).
GN PRXA
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; Core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91200671; PubMed=2016063;
RA Intapruk C., Higashimura N., Yamamoto K., Okada N., Shimmyo A.,
RA Takano M.;
RT "Nucleotide sequences of two genomic DNAs encoding peroxidase of
RT Arabidopsis thaliana.";

```

RL Gene 98:237-241(1991).
 CC -1- FUNCTION: REMOVAL OF H₂O(2), OXIDATION OF TOXIC REDUCTANTS,
 CC BIOSYNTHESIS AND DEGRADATION OF LIGNIN, DEFENSE RESPONSE TOWARD
 CC WOUNDING ON METABOLISM OF AUXIN. THESE FUNCTIONS MIGHT BE
 CC DEPENDENT ON EACH ISOZYME/ISOFORM IN EACH PLANT TISSUE.
 CC -1- CATALYTIC ACTIVITY: DONOR + H₂O(2) = OXIDIZED DONOR + 2 H₂O.
 CC -1- COFACTOR: HEME.
 CC -1- TISSUE SPECIFICITY: ROOTS.
 CC -1- SIMILARITY: BELONGS TO THE PEROXIDASE FAMILY. PLANT PEROXIDASE
 CC SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M58381; AAA32842.1; -;
 CC PIR: J00458; J00458.
 CC HSSP: P00433; LATJ.
 CC InterPro: IPR000823; -;
 CC DR InterPro: IPR002016; -;
 CC DR Pfam: PF00141; peroxidase; 1.
 CC DR PRINTS: PR00458; PEROXIDASE.
 CC DR PROSITE: PS00435; PEROXIDASE_1; 1.
 CC DR PROSITE: PS00436; PEROXIDASE_2; 1.
 CC KW Oxidoreductase; Glycoprotein; Peroxidase; Heme; Multigene family;
 KW Signal.
 FT SIGNAL 1 19
 FT CHAIN 20 349 BASIC PEROXIDASE E.
 FT ACT_SITE 67 67 BY SIMILARITY.
 FT ACT_SITE 71 71 DISTAL HISTIDINE (BY SIMILARITY).
 FT ACT_SITE 199 199 PROXIMAL HISTIDINE (HEME AXIAL LIGAND).
 FT DISULFID 40 120 BY SIMILARITY.
 FT DISULFID 73 78 BY SIMILARITY.
 FT DISULFID 126 329 BY SIMILARITY.
 FT DISULFID 206 238 BY SIMILARITY.
 FT CARBOHYD 86 86 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 173 173 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 187 187 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 217 217 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 243 243 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 349 AA; 38172 MW; 6E02C5758C36AB8F CRC64;
 Query Match 78.0%; Score 32; DB 1; Length 349;
 Best Local Similarity 83.3%; Pred. No. 26;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 2 MYPTYL 7
 DB 226 LYPTYL 231
 RESULT 5
 ID NUSM_HIPAM STANDARD; PRT; 594 AA.
 AC Q922Y1;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DE NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 5 (EC 1.6.5.3).
 GN MTND5 OR ND5 OR NADH5.
 OS Hippopotamus amphibius (Hippopotamus).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Hippopotamidae; Hippopotamus.
 OX NCBI_TaxID=9833;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Ursing B.M., Arnason U.;
 RT Analyses of mitochondrial genomes strongly support a hippopotamus-
 RT whale clade.";
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AJ010957; CAA09438.1; -;
 CC InterPro: IPR001516; -;
 CC DR InterPro: IPR001750; -;
 CC DR Pfam: PF00361; oxidored_g1; 1.
 CC DR Pfam: PF00662; oxidored_g1_N; 1.
 KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
 SQ SEQUENCE 594 AA; 66599 MW; 4130499096B5A5CE CRC64;
 Query Match 78.0%; Score 32; DB 1; Length 594;
 Best Local Similarity 83.3%; Pred. No. 44;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 3 YPTYLK 8
 DB 32 YPTVVK 37
 RESULT 6
 ID RPOB_THEMA STANDARD; PRT; 1263 AA.
 AC P29398;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE DNA-DIRECTED RNA POLYMERASE BETA CHAIN (EC 2.7.7.6) (TRANSCRIPTASE
 DE BETA CHAIN) (RNA POLYMERASE BETA SUBUNIT).
 GN RPOB OR TM0458
 OS Thermotoga maritima.
 OC Bacteria; Thermotogales; Thermotoga.
 OX NCBI_TaxID=2336;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MSB8 / DSM 3109;
 RX MEDLINE=94232816; PubMed=8177738;
 RA Palm P., Schleper C., Arnold-Ammer I., Holz I., Meier T.,
 RA Lottspeich F., Zillig W.;
 RT "The DNA-dependent RNA-polymerase of Thermotoga maritima;
 RT characterisation of the enzyme and the DNA-sequence of the genes for
 RT the large subunits.";
 RL Nucleic Acids Res. 21:4904-4908(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MSB8 / DSM 3109;
 RX MEDLINE=99287316; PubMed=10360571;
 RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
 RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
 RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
 RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
 RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
 RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
 RT "Evidence for lateral gene transfer between Archaea and Bacteria from
 RT genome sequence of Thermotoga maritima.";
 RL Nature 399:323-329(1999).
 RN [3]
 RP SEQUENCE OF 1-404 FROM N.A.
 RC STRAIN=MSB8 / DSM 3109;
 RX MEDLINE=93054590; PubMed=1429627;
 RN Liao D., Dennis P.P.

RT "The organization and expression of essential transcription
 RT translation component genes in the extremely thermophilic eubacterium
 RT Thermotoga maritima.";
 RL J. Biol. Chem. 267:22787-22797(1992).
 CC -!- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
 CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
 CC SUBSTRATES.
 CC -!- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE = N PYROPHOSPHATE +
 CC RNA(N).
 CC -!- SUBUNIT: THE ENZYME CONSISTS OF THE SIGMA CHAIN AND THE CORE
 CC ENZYME CHAIN WHICH IS COMPOSED OF 2 ALPHA CHAINS, 1 BETA CHAIN, AND 1
 CC BETA' CHAIN.
 CC -!- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA CHAIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X72695; CAA51246.1;
 CC EMBL; AE001724; AAD35543.1;
 CC EMBL; Z11839; CAA77863.1;
 CC PIR; S19903; S19903.
 CC PIR; S41466; S41466.
 CC TIGR; TM0458;
 CC InterPro; IPR001572;
 CC Pfam; PF00562; RNA_POL_B; 1.
 CC PROSITE; PS01166; RNA_POL_BETA; 1.
 KW Transcription; Transcription; DNA-directed RNA polymerase.
 SQ SEQUENCE 1263 AA; 143137 MW; 04B79368567C8237 CRC64;

Query Match 78.0%; Score 32; DB 1; Length 1263;
 Best Local Similarity 83.3%; Pred. No. 94;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYL 7
 DB 214 LYPTYL 219

RESULT 7
 MM07_FELCA STANDARD; PRT; 262 AA.
 ID MM07_FELCA
 AC P55032;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE MATRILYSIN PRECURSOR (EC 3.4.24.23) (PUMP-1 PROTEASE) (UTERINE
 DE METALLOPROTEINASE) (MATRIX METALLOPROTEINASE-7) (MMP-7) (MATRIN)
 DE (FRAGMENT).
 GN MMP7.
 OS Felis silvestris catus (Cat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
 OX NCBI_TaxID=9685;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Endometrium;
 RA Scailzo C.M., Verhage H.G., Jaffe R.C.;
 RL "Expression and estrogen control of PUMP-1 mRNA in the cat uterus.";
 RL Endocrinol. Jpn. 2:229-235(1994).
 CC -!- FUNCTION: DEGRADATES CASEIN, GELATINS OF TYPES I, III, IV, AND V,
 CC AND FIBRONECTIN. ACTIVATES PROCOLLAGENASE.
 CC -!- CATALYTIC ACTIVITY: CLEAVAGE OF 14-ALA-1-LEU-15 AND 16-TYR-1-
 CC LEU-17 IN B CHAIN OF INSULIN. NO ACTION ON COLLAGEN TYPES I, II,
 CC IV, V. CLEAVES GELATIN CHAIN ALPHA-2(I) > ALPHA-1(I).
 CC -!- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC
 CC METALLOPROTEASE) ALSO KNOWN AS MATRININ SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U04444; AAA18222.1;
 CC HSP: P09237; INMR.
 CC MEROPS; M10.008;
 CC InterPro; IPR000130;
 CC InterPro; IPR001818;
 CC Pfam; PF00413; Peptidase_M10; 1.
 CC PROSITE; PS00142; ZINC_PROTEASE; 1.
 CC PROSITE; PS00546; CYSTEINE_SWITCH; FALSE_NEG.
 KW Hydrolase; Metalloprotease; Zinc; Zymogen; Calcium;
 KW Collagen degradation; Extracellular matrix; Signal.
 FT NON_TER 1
 FT SIGNAL <1 12 BY SIMILARITY.
 FT PROPEP 13 89 ACTIVATION PEPTIDE (BY SIMILARITY).
 FT CHAIN 90 262 MATRILYSIN.
 FT SITE 82 82 CYSTEINE SWITCH (POTENTIAL).
 FT METAL 209 209 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACT_SITE 210 210 BY SIMILARITY.
 FT METAL 213 213 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 219 219 ZINC (CATALYTIC) (BY SIMILARITY).
 SQ SEQUENCE 262 AA; 29263 MW; E4A1FA23320DC732 CRC64;

Query Match 75.6%; Score 31; DB 1; Length 262;
 Best Local Similarity 100.0%; Pred. No. 31;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYL 6
 DB 227 MYPTL 231

RESULT 8
 MM07_MOUSE STANDARD; PRT; 264 AA.
 ID MM07_MOUSE
 AC Q10738;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE MATRILYSIN PRECURSOR (EC 3.4.24.23) (PUMP-1 PROTEASE) (UTERINE
 DE METALLOPROTEINASE) (MATRIX METALLOPROTEINASE-7) (MMP-7) (MATRIN).
 GN MMP7.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ICR; TISSUE=Uterus;
 RA Wilson C.B., Heppner K.J., Rudolph L.A., Matrisian L.M.;
 RL "The metalloproteinase matrilysin is preferentially expressed by
 RL epithelial cells in a tissue-restricted pattern in the mouse.";
 RL Mol. Biol. Cell 6:851-869(1995).
 CC -!- FUNCTION: DEGRADATES CASEIN, GELATINS OF TYPES I, III, IV, AND V,
 CC AND FIBRONECTIN. ACTIVATES PROCOLLAGENASE (BY SIMILARITY).
 CC -!- FUNCTION: MAY PLAY A ROLE IN TISSUE REORGANIZATION.
 CC -!- CATALYTIC ACTIVITY: CLEAVAGE OF 14-ALA-1-LEU-15 AND 16-TYR-1-
 CC LEU-17 IN B CHAIN OF INSULIN. NO ACTION ON COLLAGEN TYPES I, II,
 CC IV, V. CLEAVES GELATIN CHAIN ALPHA-2(I) > ALPHA-1(I).
 CC -!- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC
 CC METALLOPROTEASE) ALSO KNOWN AS MATRININ SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch.)

CC -----
 CC EMBL; L36238; AAA99984.1; -
 DR EMBL; L36243; AAA99984.1; JOINED.
 DR EMBL; L36242; AAA99984.1; JOINED.
 DR EMBL; L36241; AAA99984.1; JOINED.
 DR EMBL; L36240; AAA99984.1; JOINED.
 DR EMBL; L36239; AAA99984.1; JOINED.
 DR EMBL; L36244; AAA99983.1; -
 DR HSP; P09237; 1MMR.
 DR MGD; MGI:103189; Mmp7.
 DR InterPro; IPR000130; -
 DR Pfam; PF00413; Peptidase_M10; 1.
 DR PRINTS; PR00138; MATRILYSIN.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 DR PROSITE; PS00546; CYSTEINE_SWITCH; 1.
 KW Collagenase; Metalloprotease; Zinc; Zymogen; Calcium;
 FT SIGNAL 1 17 POTENTIAL.
 FT PROPEP 18 94 ACTIVATION PEPTIDE (BY SIMILARITY).
 FT CHAIN 95 264 MATRILYSIN.
 FT SITE 87 87 CYSTEINE SWITCH (POTENTIAL).
 FT METAL 214 214 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACT_SITE 215 215 BY SIMILARITY.
 FT METAL 218 218 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 224 224 ZINC (CATALYTIC) (BY SIMILARITY).
 FT CONFLICT 201 201 G -> D (IN AAA99983).
 SQ SEQUENCE 264 AA; 29755 MW; ED431A5EBAC63342 CRC64;

Query Match 75.6%; Score 31; DB 1; Length 264;

Best Local Similarity 100.0%; Pred. No. 31;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6

Db 232 MYPTY 236

RESULT 9
 MM07_HUMAN STANDARD; PRT; 267 AA.
 P09237;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE MATRILYSIN PRECURSOR (EC 3.4.24.23) (PUMP-1 PROTEASE) (UTERINE
 DE METALLOPROTEINASE) (MATRIX METALLOPROTEINASE-7) (MMP-7) (MATRIN).
 GN MP7 OR MP5L1 OR PUMP1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]

SEQUENCE FROM N.A.
 RX MEDLINE=88339885; PubMed=2844164;
 RA Muller D., Quantin B., Gesnel M.-C., Millon-Collard R., Abecassis J.,
 RA Breathnach R.;
 RT "The collagenase gene family in humans consists of at least four
 RT members.";
 RL Biochem. J. 253:187-192(1988).
 RN [2]

SEQUENCE FROM N.A.

TISSUE=Kidney;

MEDLINE=92359961; PubMed=1497627;

Marti H.P., McNeil L., Thomas G., Davies M., Lovett D.H.;

RT "Molecular characterization of a low-molecular-mass matrix
 RT metalloproteinase secreted by glomerular mesangial cells as PUMP-1.";
 RL Biochem. J. 285:899-905(1992).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=94124555; PubMed=8294454;
 RA Gaire M., Magbanua Z., McDonnell S., McNeil L.B., Lovett D.H.,
 RA Matrisian L.M.;
 RT "Structure and expression of the human gene for the matrix
 RT metalloproteinase matrilysin.";
 RL J. Biol. Chem. 269:2032-2040(1994).
 RN [4]
 RP SEQUENCE OF 18-42.
 RX MEDLINE=91070531; PubMed=2253219;
 RA Miyazaki K., Hattori Y., Umenishi F., Yasumitsu H., Umeda M.;
 RT "Purification and characterization of extracellular matrix-degrading
 RT metalloproteinase, matrin (pump-1), secreted from human rectal
 RT carcinoma cell line.";
 RL Cancer Res. 50:7758-7764(1990).
 RN [5]
 RP FUNCTION.
 RX MEDLINE=89375247; PubMed=2550050;
 RA Quantin B., Murphy G., Breathnach R.;
 RT "Pump-1 cDNA codes for a protein with characteristics similar to
 RT those of classical collagenase family members.";
 RL Biochemistry 28:5327-5334(1989).
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (2.3-ANGSTROMS).
 RX MEDLINE=95275856; PubMed=7756291;
 RA Browner M.F., Smith W.W., Castelhan A.L.;
 RT "Matrilysin-inhibitor complexes: common themes among
 RT metalloproteases.";
 RL Biochemistry 34:6602-6610(1995).
 CC 1- FUNCTION: DEGRADATES CASEIN, GELATINS OF TYPES I, III, IV, AND V,
 CC AND FIBRONECTIN. ACTIVATES PROCOLLAGENASE.
 CC 1- CATALYTIC ACTIVITY: CLEAVAGE OF 14-ALA-1-LEU-15 AND 16-TR-1-
 CC LEU-17 IN B CHAIN OF INSULIN. NO ACTION ON COLLAGEN TYPES I, II,
 CC IV, V. CLEAVES GELATIN CHAIN ALPHA-2(I) > ALPHA-1(I).
 CC 1- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY.
 CC 1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC
 CC METALLOPROTEASE) ALSO KNOWN AS MATRILYSIN SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch.)

CC EMBL; X07819; CAA30678.1; -
 DR EMBL; Z11887; CAA77942.1; -
 DR EMBL; L22524; AAC37543.1; -
 DR EMBL; L22519; AAC37543.1; JOINED.
 DR EMBL; L22520; AAC37543.1; JOINED.
 DR EMBL; L22521; AAC37543.1; JOINED.
 DR EMBL; L22522; AAC37543.1; JOINED.
 DR EMBL; L22523; AAC37543.1; JOINED.
 DR PIR; B28816; KCHUM.
 DR PIR; S24324; S24324.
 DR PDB; 1MMP; 03-APR-96.
 DR PDB; 1MMQ; 03-APR-96.
 DR PDB; 1MMR; 03-APR-96.
 DR MEROPS; M10.008; -
 DR MIM; I78990; -
 DR InterPro; IPR000130; -
 DR InterPro; IPR001818; -
 DR Pfam; PF00413; Peptidase_M10; 1.
 DR PRINTS; PR00138; MATRILYSIN.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 DR PROSITE; PS00546; CYSTEINE_SWITCH; 1.
 KW Hydrolase; Metalloprotease; Zinc; Zymogen; Calcium; 3D-structure;

EMBL; L24374; AAA99432.1; -,
HSP; P09237; 1MMR.
MEROPS; M10.008; -,
InterPro; IPR000130; -,
InterPro; IPR001818; -,
Pfam; PF00413; Peptidase_M10; 1.
PRINTS; P00138; Peptidase_M10;
PROSITE; PS00142; ZINC_PROTEASE; 1.
PROSITE; PS00546; CYSTEINE_SWITCH; 1.
Hydrolase; Metalloprotease; Zinc; Zymogen; Calcium;
Collagen degradation; Extracellular matrix; Signal.

[illegible]

Query Match 75.6%; Score 31; DB 1; Length 301;
 Best Local Similarity 100.0%; Pred. No. 36;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
 DB 106 MYPTY 110

RESULT 12
 MM12_MOUSE STANDARD; PRT; 462 AA.
 ID MM12_MOUSE AC P34960;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE MACROPHAGE METALLOELASTASE PRECURSOR (EC 3.4.24.65) (MME) (MATRIX
 METALLOPROTEINASE-12) (MMP-12).
 MM12 OR MMEL OR MME.
 Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 99-125.
 RC TISSUE=Macrophage;
 RX MEDLINE=92165826; PubMed=1537850;
 RA Shapiro S.D., Griffin G.L., Gilbert D.J., Jenkins N.A.,
 RA Copeland N.G., Welgus H.G., Senior R.M., Ley T.J.;
 RT "Molecular cloning, chromosomal localization, and bacterial
 expression of a murine macrophage metalloelastase.";
 RL J. Biol. Chem. 267:4664-4671(1992).
 CC -!- FUNCTION: MAY BE INVOLVED IN TISSUE INJURY AND REMODELING. HAS
 CC SIGNIFICANT ELASTOLYTIC ACTIVITY.
 CC -!- CATALYTIC ACTIVITY: HYDROLYSIS OF SOLUBLE AND INSOLUBLE ELASTIN.
 CC SPECIFIC CLEAVAGES ARE ALSO PRODUCED AT 14-ALA-|-LEU-15 AND 16-
 CC TYR-|-LEU-17 IN THE B CHAIN OF INSULIN.
 CC -!- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY.
 CC -!- SIMILARITY: CONTAINS 1 HEMOPEXIN-LIKE DOMAIN.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC
 CC METALLOPROTEASE) ALSO KNOWN AS MATRIXIN SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M82831; AAA39526.1; -
 DR PIR; A42401; A42401.
 DR HSP; P03956; ICGL.
 DR MEROPS; M10.009; -
 DR MGD; MGI:97005; Mmp12.
 DR InterPro; IPR000130; -
 DR InterPro; IPR000585; -
 DR InterPro; IPR001818; -
 DR Pfam; PF00413; Peptidase_M10; 1.
 DR Pfam; PF00045; hemopexin; 4.
 DR PRINTS; PR00138; MATRIXIN.
 DR PROSITE; PS00024; HEMOPEXIN; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 DR PROSITE; PS00546; CYSTEINE_SWITCH; 1.
 DR Extracellular matrix; Signal.
 KW Hydrolyase; Metalloprotease; Glycoprotein; Zinc; Zymogen; Calcium;
 FT SIGNAL 1 17 PROBABLE.
 FT PROPEP 18 98 ACTIVATION PEPTIDE.
 FT CHAIN 99 462 MACROPHAGE METALLOELASTASE.
 FT DOMAIN 272 462 HEMOPEXIN-LIKE.
 FT SITE 85 85 CYSTEINE SWITCH (BY SIMILARITY).

FT METAL 211 211 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACT_SITE 212 212 BY SIMILARITY.
 FT METAL 215 215 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 221 221 ZINC (CATALYTIC) (BY SIMILARITY).
 FT CARBOHYD 21 21 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 74 74 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 310 310 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT DISULFID 275 462 BY SIMILARITY.
 SQ SEQUENCE 462 AA; 53841 MW; BB9625906FIDBEDF CRC64;
 Query Match 75.6%; Score 31; DB 1; Length 462;
 Best Local Similarity 100.0%; Pred. No. 55;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 MYPTY 6
 DB 229 MYPTY 233
 RESULT 13
 MM12_RAT STANDARD; PRT; 465 AA.
 ID MM12_RAT AC Q63341;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE MACROPHAGE METALLOELASTASE PRECURSOR (EC 3.4.24.65) (MME) (MATRIX
 METALLOPROTEINASE-12) (MMP-12).
 MM12 OR MMEL.
 GN Rattus norvegicus (Rat).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY;
 RA Cossins J., Clements J., Catlin G.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: MAY BE INVOLVED IN TISSUE INJURY AND REMODELING. HAS
 CC SIGNIFICANT ELASTOLYTIC ACTIVITY (BY SIMILARITY).
 CC -!- CATALYTIC ACTIVITY: HYDROLYSIS OF SOLUBLE AND INSOLUBLE ELASTIN.
 CC SPECIFIC CLEAVAGES ARE ALSO PRODUCED AT 14-ALA-|-LEU-15 AND 16-
 CC TYR-|-LEU-17 IN THE B CHAIN OF INSULIN.
 CC -!- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY (BY SIMILARITY).
 CC -!- SIMILARITY: CONTAINS 1 HEMOPEXIN-LIKE DOMAIN.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC
 CC METALLOPROTEASE) ALSO KNOWN AS MATRIXIN SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X98517; CAA67142.1; -
 DR HSP; P03956; ICGL.
 DR MEROPS; M10.009; -
 DR InterPro; IPR000130; -
 DR InterPro; IPR000585; -
 DR InterPro; IPR001818; -
 DR Pfam; PF00413; Peptidase_M10; 1.
 DR PRINTS; PR00138; MATRIXIN.
 DR PROSITE; PS00024; HEMOPEXIN; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 DR PROSITE; PS00546; CYSTEINE_SWITCH; 1.
 DR Hydrolyase; Metalloprotease; Glycoprotein; Zinc; Zymogen; Calcium;
 KW Extracellular matrix; Signal; PROBABLE.
 FT SIGNAL 1 21
 FT PROPEP 22 101 ACTIVATION PEPTIDE (BY SIMILARITY).

```
FT CHAIN 102 465 MACROPHAGE METALLOELASTASE.
FT SITE 275 465 HEMOPEXIN-LIKE.
FT METAL 88 88 CYSTEINE SWITCH (BY SIMILARITY).
FT ACT_SITE 214 214 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 215 215 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 218 218 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 224 224 ZINC (CATALYTIC) (BY SIMILARITY).
FT CARBOHYD 313 313 N-LINKED (GLCNAC. .) (POTENTIAL).
FT DISULFID 278 465 BY SIMILARITY.
SQ SEQUENCE 465 AA; 53738 MW; E779B6014EC6FF68 CRC64;

Query Match 75.6%; Score 31; DB 1; Length 465;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
DB 232 MYPTY 236

RESULT 14
MM18_XENLA STANDARD; PRT; 467 AA.
AC OL3065;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DE MATRIX METALLOPROTEINASE-18 PRECURSOR (EC 3.4.24.-) (MMP-18)
DE (COLLAGENASE-4) (COLLAGENASE 4) (XCOLA4).
GN MMP18 OR COL4.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC xenopodinae; xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Intestine;
RX MEDLINE=97053976; PubMed=8898355;
RA Stow M.A., Bauzon D.D., Li J., Sedgwick T., Liang V.C.-T.,
RA Sang Q.A., Shi Y.-B.;
RT Identification and characterization of a novel collagenase in Xenopus
RT laevis: possible roles during frog development.;
RL Mol. Biol. Cell 7:1471-1483(1996).
CC -!- FUNCTION: CLEAVES COLLAGEN TYPE I. MAY PLAY A ROLE IN LARVAL
CC METAMORPHOSIS. MAY BE INVOLVED IN TAIL RESORPTION.
CC -!- COPACITOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY (BY SIMILARITY).
CC -!- ENZYME REGULATION: UPREGULATED IN THE TAIL BY THYROID HORMONE.
CC -!- SUBCELLULAR LOCATION: SECRETED. ASSOCIATED WITH THE EXTRACELLULAR
CC MATRIX (BY SIMILARITY).
CC -!- TISSUE SPECIFICITY: EXPRESSED ONLY TRANSIENTLY IN WHOLE ANIMAL, AT
CC TIME WHEN TADPOLE FEEDING BEGINS.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED AT HIGH LEVEL AS THE TADPOLE TAIL
CC RESORBS AND DURING HINDLIMB MORPHOGENESIS AND INTESTINAL
CC REMODELING.
CC -!- SIMILARITY: CONTAINS 1 HEMOPEXIN-LIKE DOMAIN.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC
CC METALLOPROTEASE) ALSO KNOWN AS MATRININ SUBFAMILY.
-----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; L76275; AAB53148.1; -.
CC HSSP; P03956; 1CGL.
CC MEROPS; M10.018; -.
CC InterPro; IPR000130; .
```

```
DR InterPro; IPR000585; -.
DR Pfam; PF00045; hemopexin; 4.
DR Pfam; PF00413; Peptidase_M10; 1.
DR PRINTS; PR00138; MATRININ.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
DR PROSITE; PS00546; CYSTEINE_SWITCH; FALSE_NEG.
DR PROSITE; PS00024; HEMOPEXIN; FALSE_NEG.
KW Hydrolase; Metalloprotease; Zinc; Calcium; Zymogen; Signal;
KW Collagen degradation; Extracellular matrix.
FT SIGNAL 1 17 POTENTIAL.
FT PROPEP 18 99 BY SIMILARITY.
FT CHAIN 100 467 MATRIX METALLOPROTEINASE-18.
FT DOMAIN 277 467 HEMOPEXIN-LIKE.
FT SITE 92 92 CYSTEINE SWITCH (POTENTIAL).
FT METAL 218 218 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 219 219 BY SIMILARITY.
FT METAL 222 222 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 228 228 ZINC (CATALYTIC) (BY SIMILARITY).
FT DISULFID 280 467 POTENTIAL.
SQ SEQUENCE 467 AA; 52812 MW; 4623F6CEFF3454051 CRC64;

Query Match 75.6%; Score 31; DB 1; Length 467;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
DB 236 MYPTY 240

RESULT 15
MM20_MOUSE STANDARD; PRT; 482 AA.
AC P57748;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE MATRIX METALLOPROTEINASE-20 PRECURSOR (EC 3.4.24.-) (MMP-20) (ENAMEL
DE METALLOPROTEINASE) (ENAMELYSIN).
GN MMP20.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SVJ;
RX MEDLINE=20079167; PubMed=10610728;
RA Caterina J., Shi J., Krakora S., Bartlett J.D., Engler J.A.,
RA Kozak C.A., Birkedal-Hansen H.;
RT Isolation, characterization, and chromosomal location of the mouse
RT enamelysin gene.;
RL Genomics 62:308-311(1999).
CC -!- FUNCTION: DEGRADATES AMELOGENIN, THE MAJOR PROTEIN COMPONENT OF THE
CC ENAMEL MATRIX AND TWO OF THE MACROMOLECULES CHARACTERISING THE
CC CARTILAGE EXTRACELLULAR MATRIX: AGGREGAN AND THE CARTILAGE
CC OLIGOMERIC MATRIX PROTEIN (COMP). MAY PLAY A CENTRAL ROLE IN TOOTH
CC ENAMEL FORMATION (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: CLEAVES AGGREGAN AT THE 360-ASN-1-PHE-361 SITE
CC (BY SIMILARITY).
CC -!- COPACITOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: SECRETED. ASSOCIATED WITH THE EXTRACELLULAR
CC MATRIX (BY SIMILARITY).
CC -!- TISSUE SPECIFICITY: EXPRESSED SPECIFICALLY IN THE ENAMEL ORGAN.
CC -!- PTM: AUTOACTIVATES AT LEAST AT THE 106-ASN-1-TYR-107 SITE (BY
CC SIMILARITY).
CC -!- SIMILARITY: CONTAINS 1 HEMOPEXIN-LIKE DOMAIN.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC
CC METALLOPROTEASE) ALSO KNOWN AS MATRININ SUBFAMILY.
-----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; AFI56956; AAF28472.1; -
DR EMBL; AFI56947; AAF28472.1; JOINED.
DR EMBL; AFI56948; AAF28472.1; JOINED.
DR EMBL; AFI56949; AAF28472.1; JOINED.
DR EMBL; AFI56950; AAF28472.1; JOINED.
DR EMBL; AFI56951; AAF28472.1; JOINED.
DR EMBL; AFI56952; AAF28472.1; JOINED.
DR EMBL; AFI56953; AAF28472.1; JOINED.
DR EMBL; AFI56954; AAF28472.1; JOINED.
DR EMBL; AFI56955; AAF28472.1; JOINED.
DR EMBL; AFI56956; AAF28472.1; JOINED.
DR MEROPS; M10.019; -
DR MGD; MGI:1353466; Mmp20.
DR PROSITE; PS00546; CYSTEINE_SWITCH; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
DR PROSITE; PS00024; HEMOPEXIN; FALSE.NEG.
KW Hydrolase; Metalloprotease; zinc; Calcium; Zymogen; Signal;
KW Extracellular matrix.
FT SIGNAL 1 21 POTENTIAL.
FT PROPEP 22 106 BY SIMILARITY.
FT CHAIN 107 482 MATRIX METALLOPROTEINASE-20.
FT DOMAIN 292 482 HEMOPEXIN-LIKE.
FT SITE 99 99 CYSTEINE SWITCH (POTENTIAL).
FT METAL 225 225 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 226 226 BY SIMILARITY.
FT METAL 229 229 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 235 235 ZINC (CATALYTIC) (BY SIMILARITY).
FT DISULFID 295 482 BY SIMILARITY.
SQ SEQUENCE 482 AA; 54373 MW; 366149FAF2BDC8BB CRC64;

Query Match 75.6%; Score 31; DB 1; Length 482;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 MYPTY 6
|
|
|
|
|
Db 243 MYPTY 247

arch completed: September 4, 2001, 15:54:10
b time: 307 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2001, 15:53:36 ; Search time 80.44 Seconds
(without alignments)
13.158 Million cell updates/sec

Title: US-09-630-345-5
Perfect score: 41
Sequence: 1 XMYPTYLK 8

Scoring table: BLOSUM62
Gapop.10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL16.*
1: sp-archaea.*
2: sp-bacteria.*
3: sp-fungi.*
4: sp-human.*
5: sp-invertebrate.*
6: sp-mammal.*
7: sp-mhc.*
8: sp-organelle.*
9: sp-phage.*
10: sp-plant.*
11: sp-rodent.*
12: sp-unclassified.*
13: sp-vertebrate.*
14: sp-virus.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35	85.4	1088	4 Q9H2T7	Q9h2t7 homo sapien
2	33	80.5	269	1 Q27159	Q27159 methanobact
3	33	80.5	1089	10 Q9LIT28	Q9lit28 arabidopsis
4	32	78.0	252	2 Q9L8T5	Q9l8t5 haemophilus
5	32	78.0	553	11 Q70577	Q70577 mus musculus
6	32	78.0	555	11 Q9R0W2	Q9r0w2 rattus norv
7	32	78.0	593	11 P97558	P97558 rattus norv
8	32	78.0	593	11 P70485	P70485 rattus norv
9	32	78.0	986	4 Q94858	Q94858 homo sapien
10	32	78.0	1346	4 Q9UI23	Q9uiz3 homo sapien
11	31	75.6	40	14 Q9I2W8	Q9i2w8 human coxa
12	31	75.6	40	14 Q9I2W7	Q9i2w7 human coxa
13	31	75.6	50	14 Q66796	Q66796 human echov
14	31	75.6	53	6 Q9XS83	Q9xs83 equus cabal
15	31	75.6	198	14 Q9Q907	Q9q907 shope fibro
16	31	75.6	261	4 Q9NR81	Q9nr81 homo sapien
17	31	75.6	261	4 Q9NR87	Q9nr87 homo sapien
18	31	75.6	261	4 Q9GZS2	Q9gzs2 homo sapien
19	31	75.6	262	5 Q19918	Q19918 caenorhabdi

20	31	75.6	267	6 Q9TV55	Q9tv55 sus scrofa
21	31	75.6	291	5 Q94467	Q94467 dictyosteli
22	31	75.6	307	5 Q9VJ65	Q9vj65 drosophila
23	31	75.6	332	2 Q51888	Q51888 prevotella
24	31	75.6	378	10 Q9SHH5	Q9shh5 arabidopsis
25	31	75.6	385	10 Q9SKA2	Q9ska2 arabidopsis
26	31	75.6	420	4 Q9H4D2	Q9hd2 homo sapien
27	31	75.6	445	5 Q9VPA2	Q9vpa2 drosophila
28	31	75.6	483	13 Q98857	Q98857 cynops pyrr
29	31	75.6	497	14 Q96640	Q96640 bovine papi
30	31	75.6	502	5 Q18426	Q18426 caenorhabdi
31	31	75.6	516	5 Q18198	Q18198 caenorhabdi
32	31	75.6	565	1 Q29350	Q29350 archaeoglob
33	31	75.6	585	5 Q9TXN1	Q9txn1 caenorhabdi
34	31	75.6	769	11 Q9ESN3	Q9esn3 mus musculu
35	31	75.6	2185	14 Q88445	Q88445 swine vesic
36	30	73.2	137	14 Q9J919	Q9j919 human coxa
37	30	73.2	141	5 Q02173	Q02173 caenorhabdi
38	30	73.2	152	2 Q9I401	Q9i401 pseudomonas
39	30	73.2	155	14 Q98534	Q98534 paramecium
40	30	73.2	158	5 Q9VUZ1	Q9vuz1 drosophila
41	30	73.2	188	5 Q9V5Y8	Q9v5y8 drosophila
42	30	73.2	238	14 Q9QB94	Q9qb94 yaba monkey
43	30	73.2	248	2 Q48871	Q48871 lactobacill
44	30	73.2	255	14 Q9DHM6	Q9dhm6 yaba-like d
45	30	73.2	262	2 Q9PIA0	Q9pia0 campylobact

ALIGNMENTS

RESULT 1

Q9H2T7 ID Q9H2T7 PRELIMINARY; PRT; 1088 AA.
AC Q9H2T7;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE RANBP17.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=20568271; PubMed=11024021;

RA Kutay U., Hartmann E., Treichel N., Calado A., Carmo-Fonseca M.,

RA Prehn S., Kraft R., Gorlich D., Bischoff F.R.;

RT Identification of Two Novel RanGTP-binding Proteins Belonging to the

RT Importin beta Superfamily.;

RL J. Biol. Chem. 275:40163-40168(2000).

DR EMBL; AF222747; AAG44255.1; -

SQ SEQUENCE 1088 AA; 124374 MW; A95D559388EAEBF CRC64;

Query Match 85.4%; Score 35; DB 4; Length 1088;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYL 7

|||||

Db 744 MYPTYL 749

RESULT 2

ID Q27159 PRELIMINARY; PRT; 269 AA.
AC Q27159;

DT 01-JAN-1998 (TREMBLrel. 05, Created)

DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)

DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)

DE HYPOTHETICAL 31.5 KDA PROTEIN.

GN MTH1087.

Qy 2 MYPTYL 7
||||:

Query Match 78.0%; Score 32; DB 11; Length 553;

Best Local Similarity 57.1%; Pred. No. 1.8e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 4; Conservative 3;

QY 2 MYPTYLK 8
:||||:
Db 457 LYPTIIR 463

RESULT 6
Q9ROW2 PRELIMINARY; PRT; 555 AA.
AC Q9ROW2
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE ORGANIC CATION TRANSPORTER OCT2R.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
[1]
SEQUENCE FROM N.A.
RC TISSUE=KIDNEY;
RX MEDLINE=99316020; PubMed=10385678;
RA Gruendemann D., Liebich G., Kiefer N., Koester S., Schoemig E.;
RT "Selective substrates for non-neuronal monoamine transporters.";
RL Mol. Pharmacol. 56:1-10(1999).
DR EMBL; Y13154; CAB52215.1; -;
DR InterPro; IPR001066; -;
DR Pfam; PF00083; sugar_tr; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_2.
SQ SEQUENCE 555 AA; 62342 MW; 29521969AEIAC206 CRC64;

Query Match 78.0%; Score 32; DB 11; Length 555;
Best Local Similarity 57.1%; Pred. No. 1.8e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
:||||:
Db 457 LYPTIIR 463

RESULT 7
P97558 PRELIMINARY; PRT; 593 AA.
ID P97558
AC P97558;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE ORGANIC CATION TRANSPORTER.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
[1]
SEQUENCE FROM N.A.
RC TISSUE=KIDNEY;
RA Gorboulev V.G., Koepsell H.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
DR EMBL; X98334; CAA66979.1; -;
DR InterPro; IPR001066; -;
DR Pfam; PF00083; sugar_tr; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_2.
KW Ion transport; Transmembrane.
SQ SEQUENCE 593 AA; 66100 MW; 36C3E1B5DC057790 CRC64;

Query Match 78.0%; Score 32; DB 11; Length 593;
Best Local Similarity 57.1%; Pred. No. 2e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 4; Conservative 3;

QY 2 MYPTYLK 8
:||||:
Db 457 LYPTIIR 463

RESULT 8
P70485 PRELIMINARY; PRT; 593 AA.
ID P70485
AC P70485;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE ORGANIC CATION TRANSPORTER OCT2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
[1]
SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=KIDNEY;
RX MEDLINE=96295517; PubMed=8702418;
RA Okuda M., Saito H., Urakami Y., Takano M., Inui K.;
RT "cDNA cloning and functional expression of a novel rat kidney organic
cation transporter, OCT2";
RL Biochem. Biophys. Res. Commun. 224:500-507(1996).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
DR EMBL; D83044; BAA11754.1; -;
DR InterPro; IPR001066; -;
DR Pfam; PF00083; sugar_tr; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_2.
KW Transmembrane.
SQ SEQUENCE 593 AA; 66080 MW; 36C1044E0C04B443 CRC64;

Query Match 78.0%; Score 32; DB 11; Length 593;
Best Local Similarity 57.1%; Pred. No. 2e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
:||||:
Db 457 LYPTIIR 463

RESULT 9
O94858 PRELIMINARY; PRT; 986 AA.
ID O94858
AC O94858;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE KIAA0758 PROTEIN (FRAGMENT).
GN KIAA0758
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=99087487; PubMed=9872452;
RA Nagase T., Ishikawa K., Suyama M., Kikuno R., Miyajima N., Tanaka A.,
RA Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XI.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro";
RL DNA Res. 5:277-286(1998).
DR EMBL; AB018301; BAA34478.1; -;
DR InterPro; IPR000203; -;
DR InterPro; IPR000832; -;
DR Pfam; PF00002; 7tm_2; 1.
DR Pfam; PF01825; GPS; 1.

DR PRINTS: PR00249; GPCRSECRETIN.
DR SMART; SM00303; GPS; 1.
FT NON_TER 1
SQ SEQUENCE 986 AA; 108720 MW; EF4875B03BAA0E7D CRC64;

Query Match 78.0%; Score 32; DB 4; Length 986;
Best Local Similarity 85.7%; Pred. No. 3.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
| | | | |
DB 384 MLPTYLK 390

RESULT 10
Q9UIZ3 PRELIMINARY; PRT: 1346 AA.
AC Q9UIZ3.
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DE 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE DJ365012.1 (KIAA0758 PROTEIN).
GN DJ365012.1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Sehra H.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX DOMAIN.
DR EMBL: AL096772; CAB61578.1; -.
DR InterPro; IPR000082; -.
DR InterPro; IPR00203; -.
DR InterPro; IPR000561; -.
DR InterPro; IPR000832; -.
DR InterPro; IPR003006; -.
DR Pfam; PF00002; 7tm_2; 1.
DR Pfam; PF00047; 1g; 2.
DR Pfam; PF01825; GPS; 1.
DR PRINTS: PR00249; GPCRSECRETIN.
DR PROSITE: PS01186; EGF_2; UNKNOWN_1.
DR SMART; SM00200; SEA; 1.
SQ SEQUENCE 1346 AA; 149456 MW; 72A9D02B08218A60 CRC64;

Query Match 78.0%; Score 32; DB 4; Length 1346;
Best Local Similarity 85.7%; Pred. No. 4.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
| | | | |
DB 744 MLPTYLK 750

RESULT 11
Q9IWZ8 PRELIMINARY; PRT: 40 AA.
AC Q9IWZ8.
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE VP2 PROTEIN (FRAGMENT).
GN VP2.
OS Human coxsackievirus B5.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCBI_TaxID=12074;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN-R8597;
RA Casas I., Palacios G.F., Trallero G., Cisterna D., Freire M.C.,
RA Tenorio A.;
RT "Molecular characterization of human enteroviruses in clinical samples
by three different RT nested PCR assays and direct sequencing of
RT amplified products.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF252193; AAF71330.1; -.
DR InterPro; IPR001676; -.
DR Pfam; PF00073; rhv; 1.
DR NON_TER 1
FT NON_TER 40
SQ SEQUENCE 40 AA; 4388 MW; 7F998A9865EF9F67 CRC64;

Query Match 75.6%; Score 31; DB 14; Length 40;
Best Local Similarity 71.4%; Pred. No. 17;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
| | | | |
DB 26 VMPTYLK 32

RESULT 12
Q9IWZ7 PRELIMINARY; PRT: 40 AA.
AC Q9IWZ7.
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE VP2 PROTEIN (FRAGMENT).
GN VP2.
OS Human coxsackievirus B5.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCBI_TaxID=12074;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-R8497;
RA Casas I., Palacios G.F., Trallero G., Cisterna D., Freire M.C.,
RA Tenorio A.;
RT "Molecular characterization of human enteroviruses in clinical samples
by three different RT nested PCR assays and direct sequencing of
RT amplified products.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF252194; AAF71331.1; -.
DR InterPro; IPR001676; -.
DR Pfam; PF00073; rhv; 1.
DR NON_TER 1
FT NON_TER 40
SQ SEQUENCE 40 AA; 4388 MW; 7F998A9865EF9F67 CRC64;

Query Match 75.6%; Score 31; DB 14; Length 40;
Best Local Similarity 71.4%; Pred. No. 17;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
| | | | |
DB 26 VMPTYLK 32

RESULT 13
Q66796 PRELIMINARY; PRT: 50 AA.
ID Q66796.
AC Q66796.
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE CAPSID PROTEIN VP2 (FRAGMENT).
GN VP2.
OS Human echovirus 20.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Enterovirus.
 OX NCBI_TaxID=47508;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JV-1;
 RX MEDLINE=96203959; PubMed=8627260;
 RA Huttunen P., Santti J., Pulli T., Hyytiä T.;
 RT "The major echovirus group is genetically coherent and related to
 RT coxsackie B viruses."
 RL J. Gen. Virol. 77:715-725(1996).
 DR EMBL: X89549; CA61727.1; -;
 DR InterPro: IPR001676; -;
 DR Pfam: PF00073; rnv; 1.
 FT NON_TER 1
 FT NON_TER 50
 SQ SEQUENCE 50 AA; 5448 MW; BBFC58B2511FC2E7 CRC64;

Query Match 75.6%; Score 31; DB 14; Length 50;
 Best Local Similarity 71.4%; Pred. No. 22;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
 Db :|||||
 31 VWPYTLK 37

RESULT 14
 O9XS83 PRELIMINARY; PRT; 53 AA.
 ID Q9XS83
 AC Q9XS83;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE TUBBY PROTEIN (FRAGMENT).
 GN TUB.
 OS Equus caballus (Horse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
 OX NCBI_TaxID=9796;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99160468; PubMed=10051323;
 RA Caetano A.R., Pomp D., Murray J.D., Bowling A.T.;
 RT "Comparative mapping of 18 equine type I genes assigned by somatic
 RT cell hybrid analysis."
 RL Mamm. Genome 10:271-276(1999).
 DR EMBL: AF097580; AAD25983.1; -;
 DR HSSP: P50586; 1C8Z.
 DR InterPro: IPR000007; -;
 DR Pfam: PF01167; Tub; 1.
 FT NON_TER 1
 FT NON_TER 53
 SQ SEQUENCE 53 AA; 6244 MW; 13A9ED8C69500D3E CRC64;

Query Match 75.6%; Score 31; DB 6; Length 53;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
 Db :|||||
 16 MYPTY 20

RESULT 15
 Q9Q907 PRELIMINARY; PRT; 198 AA.
 ID Q9Q907
 AC Q9Q907;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)

DE GP064R.
 GN S064R.
 OS Shope fibroma virus (strain Kasza) (SFV).
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
 OC Leporipoxvirus.
 OX NCBI_TaxID=10272;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KASZA;
 RX MEDLINE=84165064; PubMed=6323741;
 RA Delange A.M., Macaulay C., Block W., Mueller T., McFadden G.;
 RT "Tumorigenic poxviruses: construction of the composite physical map of
 RT the Shope fibroma virus genome."
 RL J. Virol. 50:408-416(1984).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KASZA;
 RX MEDLINE=92074222; PubMed=1660196;
 RA Strayer D.S., Jerling H.H., O'Connor K.;
 RT "Sequence and analysis of a portion of the genomes of Shope fibroma
 RT virus and malignant rabbit fibroma virus that is important for viral
 RT replication in lymphocytes."
 RL Virology 185:585-595(1991).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KASZA;
 RX MEDLINE=20032074; PubMed=10562495;
 RA Willer D.O., McFadden G., Evans D.H.;
 RT "The complete genome sequence of shope (Rabbit) fibroma virus."
 RL Virology 264:319-343(1999).
 DR EMBL: AF170722; AAF17946.1; -;
 SQ SEQUENCE 198 AA; 23282 MW; 9D181052C1ED42C5 CRC64;

Query Match 75.6%; Score 31; DB 14; Length 198;
 Best Local Similarity 83.3%; Pred. No. 96;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYL 7
 Db :|||||
 110 MYPTFL 115

Search completed: September 4, 2001, 15:53:36
 Job time: 289 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2001, 15:50:25 ; Search time 77.49 Seconds
(without alignments)
6.259 Million cell updates/sec

Title: US-09-630-345-6
Perfect score: 41
Sequence: 1 XMYPTYLK 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

al number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_0601.*
1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	40	97.6	8 19 AAW50937	Vasoactive intesti
2	40	97.6	8 21 AAB08298	Vasoactive intesti
3	40	97.6	8 21 AAB08306	Amino acid sequenc
4	35	85.4	8 21 AAB08304	Amino acid sequenc
5	35	85.4	8 21 AAB08305	Amino acid sequenc
6	35	85.4	450 21 AAB63104	Human secreted pro
7	32	78.0	807 21 AAY57288	Human GPCR protein
8	32	78.0	986 21 AAY40440	Human brain-derive
9	31	75.6	8 20 AAW75457	Mammalian tub prot
10	31	75.6	12 20 AAW97073	Peptidomimetic cap
11	31	75.6	58 20 AAY12222	Human 5' EST secre

12	31	75.6	173	22	AAW74618	Human 72 kDa gelat
13	31	75.6	173	22	AAW68882	Matrilysin catalyt
14	31	75.6	261	19	AAW76253	Human matrilysin-I
15	31	75.6	264	19	AAW76254	Human matrilysin p
16	31	75.6	271	17	AAW96211	Human recombinant
17	31	75.6	297	21	AAW43585	Human cancer assoc
18	31	75.6	325	14	AAW42750	ced-4 gene II prod
19	31	75.6	459	18	AAW36485	Mouse TUB Form I.
20	31	75.6	459	21	AAW26400	Mouse tub Form I p
21	31	75.6	460	18	AAW36488	Human TUB Form 6.
22	31	75.6	460	21	AAW26903	Human TUB Form 6 p
23	31	75.6	462	22	AAW49983	Murine macrophage
24	31	75.6	469	21	AAW10655	BpV1 L1 fusion pro
25	31	75.6	469	21	AAW10660	BPV2 L1 fusion pro
26	31	75.6	483	20	AAW93163	Sequence of bovine
27	31	75.6	495	6	AAW51099	Sequence of bovine
28	31	75.6	495	20	AAW88482	Bovine papillomavi
29	31	75.6	505	18	AAW36486	Mouse TUB Form II.
30	31	75.6	505	18	AAW10728	Mouse tub gene pro
31	31	75.6	505	19	AAW54367	Mouse tub polypept
32	31	75.6	505	20	AAW75450	Mouse wild type tu
33	31	75.6	505	21	AAW26901	Mouse tub Form II
34	31	75.6	506	18	AAW36497	Human TUB Form 4.
35	31	75.6	506	18	AAW10729	Human tub gene pro
36	31	75.6	506	19	AAW54368	Human tub polypept
37	31	75.6	506	20	AAW75451	Human wild type tu
38	31	75.6	506	21	AAW26910	Human TUB form 4 5
39	31	75.6	512	18	AAW36496	Human TUB Form 3.
40	31	75.6	512	21	AAW26912	Human TUB form 3 5
41	31	75.6	518	18	AAW36495	Human TUB Form 2.
42	31	75.6	518	21	AAW26909	Human TUB form 2 5
43	31	75.6	561	18	AAW36489	Human TUB Form 1.
44	31	75.6	561	21	AAW26904	Human TUB Form 1 p
45	30	73.2	73	14	AAW40041	Fillistata peptide

ALIGNMENTS

RESULT 1
AAW50937
ID AAW50937 standard; peptide; 8 AA.
XX
AC AAW50937;
XX
DT 31-JUL-1998 (first entry)
XX
DE Vasoactive intestinal peptide receptor binding inhibitor (VIP2).
XX
KW Vasoactive intestinal peptide; VIP; antagonist; somatostatin; bombesin;
KW Substance P; cancer; inhibition.
XX
OS Synthetic.
XX
PN EP835662-A2.
XX
PD 15-APR-1998.
XX
PE 11-DEC-1996; 96EP-0309012.
XX
PR 08-OCT-1996; 96US-0727679.
PR 16-AUG-1996; 96IN-0001822.
XX
(NAIN-) NAT INST IMMUNOLOGY.
XX
Jaggi M, Mukherjee R;
WPI; 1998-208959/19.
Composition containing analogues of vasoactive intestinal peptide,
somatostatin - bombesin and substance P, for treatment of tumours
PT and for inhibiting over-expression of these peptide(s)
PT
XX

PS Claim 1; Page 4; 49pp; English.

XX The invention relates to a new composition which comprises: (i) the
CC somatostatin analogue SOM2 AGCKNFRdWKPTSDC (3-14 disulphide bridge),
CC and (ii) at least 4 of the peptides: antagonist of vasoactive
CC intestinal peptide (VIP1); VIP receptor-binding inhibitor (VIP2); VIP
CC receptor antagonist (VIP3); somatostatin analogue (SOM1); bombesin
CC antagonist (BOM1) and substance P antagonist (SP1). Also claimed are
CC more general compositions containing peptide analogues of somatostatin,
CC VIP, bombesin and substance P. The compositions are used in human or
CC veterinary medicine: (a) to kill (or inhibit multiplication of) tumour
CC or cancer cells, particularly for treatment of leukaemia, lymphoma,
CC adenocarcinoma of stomach, pancreas or prostate, or cancer of lung,
CC breast, kidney or particularly rectum and colon, and (b) to prevent,
CC inhibit or modulate over-expression of, e.g. VIP. A wide range of cancer
CC cells express receptors for VIP, somatostatin, bombesin and/or substance
CC P. The present sequence represents VIP receptor-binding inhibitor
CC (VIP2).

XX Sequence 8 AA;

Query Match 97.6%; Score 40; DB 19; Length 8;
Best Local Similarity 100.0%; Pred. No. 3.4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
DB 2 myptylk 8

RESULT 2

AAB08298
ID AAB08298 standard; peptide; 8 AA.

AC AAB08298;

DT 04-DEC-2000 (first entry)

DE Vasoactive intestinal peptide (VIP) analogue VIP2.

KW Vasoactive intestinal peptide; VIP; analogue; somatostatin; SOM1; SOM2;
KW VIP1; VIP2; VIP3; BOM1; bombesin; SP1; substance P; MuJ-7; tumour growth;
KW tumour angiogenesis; metastasis; cancer; angiogenesis; adenocarcinoma;
KW leukaemia; lymphoma.

OS Synthetic.

WO200047221-A1.

PD 17-AUG-2000.

PF 11-FEB-2000; 2000WO-US03559.

PR 11-FEB-1999; 99US-0248381.

PA (NAIM-) NAT INST IMMUNOLOGY.

PA (DABU-) DABUR RES FOUND.

PA (CORD/) CORD J I.

PI Mukherjee R, Jaggi M, Prasad S, Burman AC, Rajendran P, Mathur A;

PI Singh AT;

DR WPI; 2000-549083/50.

XX Novel therapeutically active composition comprising at least 5

XX peptides, useful for treating angiogenesis especially as a result of

XX adenocarcinomas -

XX Disclosure; Page 8; 42pp; English.

XX The present sequence represents an analogue of vasoactive intestinal

XX peptide (VIP). The specification describes therapeutically active

CC compositions comprising at least one analogue of somatostatin (chosen
CC from SOM1 and SOM2), and at least four analogues chosen from VIP1 (a
CC VIP antagonist), VIP2 (a VIP receptor binding inhibitor), VIP3 (a VIP
CC receptor antagonist), BOM1 (a bombesin antagonist), and SP1 (a substance
CC P antagonist). The combination of these 7 analogues is known as MuJ-7.
CC MuJ-7 is used as an anticancer drug to restrict tumour growth and spread
CC by inhibiting tumour angiogenesis. MuJ-7, in addition, inhibits
CC metastasis through its antiangiogenic activity in all cancers. The
CC peptides are useful for the treatment and prevention of angiogenesis,
CC especially as a result of adenocarcinomas of the colon, breast, lung,
CC prostate, kidney, leukemias or lymphomas.

SQ Sequence 8 AA;

Query Match 97.6%; Score 40; DB 21; Length 8;
Best Local Similarity 100.0%; Pred. No. 3.4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
DB 2 myptylk 8

RESULT 3

AAB08306

ID AAB08306 standard; peptide; 8 AA.

AC AAB08306;

DT 04-DEC-2000 (first entry)

DE Amino acid sequence of an antiangiogenic peptide.

KW Vasoactive intestinal peptide; VIP; analogue; somatostatin; SOM1; SOM2;
KW VIP1; VIP2; VIP3; BOM1; bombesin; SP1; substance P; MuJ-7; tumour growth;
KW tumour angiogenesis; metastasis; cancer; angiogenesis; adenocarcinoma;
KW leukaemia; lymphoma.

OS Synthetic.

Key Location/Qualifiers

FT Misc-difference 1 /note= "Optionally D-form residue"

FT Misc-difference 3 /note= "Optionally D-form residue"

FT Misc-difference 6 /note= "Optionally D-form residue"

FT Misc-difference 7 /note= "Optionally D-form residue"

FT Misc-difference 7 /note= "Optionally D-form residue"

PN WO200047221-A1.

PD 17-AUG-2000.

PF 11-FEB-2000; 2000WO-US03559.

PR 11-FEB-1999; 99US-0248381.

PA (NAIM-) NAT INST IMMUNOLOGY.

PA (DABU-) DABUR RES FOUND.

PA (CORD/) CORD J I.

PI Mukherjee R, Jaggi M, Prasad S, Burman AC, Rajendran P, Mathur A;

PI Singh AT;

DR WPI; 2000-549083/50.

XX Novel therapeutically active composition comprising at least 5

XX peptides, useful for treating angiogenesis especially as a result of

XX adenocarcinomas -

XX Claim 11; Page 31; 42pp; English.

XX AAB08304-15 represent peptides which have an antiangiogenic effect. The
 CC specification describes therapeutically active compositions comprising
 CC at least one analogue of somatostatin (chosen from SOM1 and SOM2), and
 CC at least four analogues chosen from vasoactive intestinal peptide (VIP)
 CC 1 (a VIP antagonist), VIP2 (a VIP receptor binding inhibitor), VIP3 (a
 CC VIP receptor antagonist), BOM1 (a bombesin antagonist), and SP1 (a
 CC substance P antagonist). The combination of these 7 analogues is known as
 CC MuJ-7. MuJ-7 is used as an anticancer drug to restrict tumour growth and
 CC spread by inhibiting tumour angiogenesis. MuJ-7, in addition, inhibits
 CC metastasis through its antiangiogenic activity in all cancers. The
 CC peptides are useful for the treatment and prevention of angiogenesis,
 CC especially as a result of adenocarcinomas of the colon, breast, lung,
 CC prostate, kidney, leukemias or lymphomas.
 XX
 SQ Sequence 8 AA;

Query Match 97.6%; Score 40; DB 21; Length 8;
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 MYPTYLK 8
 |||||
 Db 2 myptylk 8

RESULT 4
 AAB08304
 ID AAB08304 standard; peptide; 8 AA.
 AC AAB08304;
 XX
 DT 04-DEC-2000 (first entry)
 XX
 DE Amino acid sequence of an antiangiogenic peptide.
 XX
 KW Vasoactive intestinal peptide; VIP; analogue; somatostatin; SOM1; SOM2;
 KW VIP1; VIP2; VIP3; BOM1; bombesin; SP1; substance P; MuJ-7; tumour growth;
 KW tumour angiogenesis; metastasis; cancer; angiogenesis; adenocarcinoma;
 KW leukaemia; lymphoma.
 XX
 OS Synthetic.
 XX

Key Location/Qualifiers
 FT Modified-site 1 /label= Alb
 FT /note= "alpha-aminoisobutyric acid"
 FT Modified-site 7 /label= Alb, Deg
 FT /note= "alpha-aminoisobutyric acid or alpha, alpha
 FT diethyl glycine"
 FT
 PN WO200047221-A1.
 XX
 XX 17-AUG-2000.
 XX
 XX 11-FEB-2000; 2000WO-US03559.
 XX
 XX 11-FEB-1999; 99US-0248381.
 XX (NAIM-) NAT INST IMMUNOLOGY.
 PA (DABU-) DABUR RES FOUND.
 PA (CORD/) CORD J I.
 XX
 PI Mukherjee R, Jaggi M, Prasad S, Burman AC, Rajendran P, Mathur A;
 PI Singh AT;
 XX
 XX WPI; 2000-549083/50.
 XX
 XX Novel therapeutically active composition comprising at least 5
 PT peptides, useful for treating angiogenesis especially as a result of
 PT adenocarcinomas -

XX Claim 18; Page 36; 42pp; English.
 PS
 CC AAB08304-15 represent peptides which have an antiangiogenic effect. The
 CC specification describes therapeutically active compositions comprising
 CC at least one analogue of somatostatin (chosen from SOM1 and SOM2), and
 CC at least four analogues chosen from vasoactive intestinal peptide (VIP)
 CC 1 (a VIP antagonist), VIP2 (a VIP receptor binding inhibitor), VIP3 (a
 CC VIP receptor antagonist), BOM1 (a bombesin antagonist), and SP1 (a
 CC substance P antagonist). The combination of these 7 analogues is known as
 CC MuJ-7. MuJ-7 is used as an anticancer drug to restrict tumour growth and
 CC spread by inhibiting tumour angiogenesis. MuJ-7, in addition, inhibits
 CC metastasis through its antiangiogenic activity in all cancers. The
 CC peptides are useful for the treatment and prevention of angiogenesis,
 CC especially as a result of adenocarcinomas of the colon, breast, lung,
 CC prostate, kidney, leukemias or lymphomas.
 XX
 SQ Sequence 8 AA;

Query Match 85.4%; Score 35; DB 21; Length 8;
 Best Local Similarity 85.7%; Pred. No. 3.4e+05;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 MYPTYLK 8
 |||||
 Db 2 myptyxk 8

RESULT 5
 AAB08305
 ID AAB08305 standard; peptide; 8 AA.
 XX
 AC AAB08305;
 XX
 DT 04-DEC-2000 (first entry)
 XX

Amino acid sequence of an antiangiogenic peptide.
 XX
 DE Vasoactive intestinal peptide; VIP; analogue; somatostatin; SOM1; SOM2;
 KW VIP1; VIP2; VIP3; BOM1; bombesin; SP1; substance P; MuJ-7; tumour growth;
 KW tumour angiogenesis; metastasis; cancer; angiogenesis; adenocarcinoma;
 KW leukaemia; lymphoma.
 XX

Synthetic.
 XX
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 1 /note= "D-form residue"
 FT Modified-site 7 /label= Alb, Deg
 FT /note= "alpha-aminoisobutyric acid or alpha, alpha
 FT diethyl glycine"
 FT
 PN WO200047221-A1.
 XX
 XX 17-AUG-2000.
 XX
 XX 11-FEB-2000; 2000WO-US03559.
 XX
 XX 11-FEB-1999; 99US-0248381.
 XX (NAIM-) NAT INST IMMUNOLOGY.
 PA (DABU-) DABUR RES FOUND.
 PA (CORD/) CORD J I.
 XX
 PI Mukherjee R, Jaggi M, Prasad S, Burman AC, Rajendran P, Mathur A;
 PI Singh AT;
 XX
 XX WPI; 2000-549083/50.
 XX
 XX Novel therapeutically active composition comprising at least 5
 PT peptides, useful for treating angiogenesis especially as a result of

PT adenocarcinomas -

XX Claim 11; Page 31; 42pp; English.

XX AAB08304-15 represent peptides which have an antiangiogenic effect. The

CC specification describes therapeutically active compositions comprising

CC at least one analogue of somatostatin (chosen from SOM1 and SOM2), and

CC at least four analogues chosen from vasoactive intestinal peptide (VIP)

CC 1 (a VIP antagonist), VIP2 (a VIP receptor binding inhibitor), VIP3 (a

CC VIP receptor antagonist), BOM1 (a bombesin antagonist), and S1 (a

CC substance P antagonist). The combination of these 7 analogues is known as

CC MuJ-7. MuJ-7 is used as an anticancer drug to restrict tumour growth and

CC spread by inhibiting tumour angiogenesis. MuJ-7, in addition, inhibits

CC metastasis through its antiangiogenic activity in all cancers. The

CC peptides are useful for the treatment and prevention of angiogenesis,

CC especially as a result of adenocarcinomas of the colon, breast, lung,

CC prostate, kidney, leukemias or lymphomas.

XX Sequence 8 AA;

XX

Query Match 85.4%; Score 35; DB 21; Length 8;

Best Local Similarity 85.7%; Pred. No. 3.4e+05;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MYPTYLK 8

Db 2 myptyxk 8

|||||

RESULT 6

AAB63104

ID AAB63104 standard; Protein: 450 AA.

AC AAB63104;

XX

DT 26-MAR-2001 (first entry)

DE Human secreted protein sequence encoded by gene 18 SEQ ID NO:114.

XX

XX Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;

KW antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;

KW cerebroprotective; neurotropic; neuroprotective; antibacterial; virucide;

KW fungicide; ophthalmological; vulneprotective; gene therapy; neoplasia;

KW autoimmune disease; rheumatoid arthritis; hyperproliferative disorder;

KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;

KW cerebral ischaemia; angiogenesis; nervous system disorder; infection;

KW Alzheimer's disease; ocular disorder; corneal infection; wound healing;

XX skin aging; food additive; preservative.

OS Homo sapiens.

XX

XX WO2000061748-A1.

XX

XX 19-OCT-2000.

XX

XX 06-APR-2000; 2000WO-US08982.

XX

XX 09-APR-1999; 99US-0128696.

XX

XX 14-JAN-2000; 2000US-0176069.

XX

XX (HUMA-) HUMAN GENOME SCI INC.

XX

XX Rosen CA, Ruben SM, Komatsoulis G;

XX

XX WPI; 2000-638566/61.

XX

XX New nucleic acid molecules encoding 48 human secreted proteins for

XX diagnosing, preventing, treating or ameliorating medical conditions and

XX used as food additives or preservatives -

XX

XX Disclosure; Page 453-455; 480pp; English.

XX

Query Match 85.4%; Score 35; DB 21; Length 450;

Best Local Similarity 100.0%; Pred. No. 40;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYL 7

Db 106 myptyl 111

|||||

RESULT 7

AAV57288

ID AAV57288 standard; Protein: 807 AA.

XX

AC AAV57288;

XX

DT 05-JUN-2000 (first entry)

XX

DE Human GPCR protein (HGRP) sequence (clone ID 3036563).

XX

XX Human; G protein coupled protein receptor; HGRP; cell proliferation;

KW neurological; immune disorder; cytostatic; anti-arteriosclerotic;

KW anti-atherosclerotic; hepatotropic; antinflammatory; virucide; leukemia;

KW immunomodulatory; anemia; asthma; gastrointestinal; anti-epileptic;

KW anti-Alzheimer's; anti-Parkinsonian; gene therapy.

XX

OS Homo sapiens.

XX

XX WO2000015793-A2.

XX

XX 23-MAR-2000.

XX

XX 17-SEP-1999; 99WO-US20958.

XX

XX 17-SEP-1998; 98US-0156513.

XX

XX (INCY-) INCYTE PHARM INC.

XX

XX Bandman O, Lal P, Tang YP, Corley NC, Guegler KJ, Gorgone GA;

XX Baughn MR;

XX

XX WPI; 2000-271432/23.

XX

XX N-PSDB; AA290526.

XX

Sequence 450 AA;

PT Human G protein coupled protein receptor peptides useful for the
 PT prevention, diagnosis and treatment of cell proliferative, neurological
 PT and immune disorders -

PS Claim 1; Page 65-67; 71pp; English.

XX
 XX The invention provides human G protein coupled protein receptor (HGPRP)
 CC polypeptides and polynucleotides encoding them. The polypeptides can be
 CC produced by standard recombinant methodology. The polynucleotides and
 CC polypeptides may be used in the prevention, treatment and diagnosis of
 CC diseases associated with their inappropriate expression. Diseases that
 CC can be treated are cell proliferative disorders (e.g. arteriosclerosis,
 CC atherosclerosis and hepatitis), cancers (e.g. leukemia, melanomas and
 CC adenocarcinoma), immune disorders (e.g. anemia, asthma and Crohn's
 CC disease) and neurological disorders (e.g. epilepsy, Alzheimer's disease
 CC and Parkinson's disease). The anti-HGPRP antibodies may also be used as
 CC diagnostic agents for detecting the presence of HGPRP polypeptides in
 CC samples (e.g. by enzyme linked immunosorbent assay (ELISA)). Sequences
 CC AAY57283-288 represent the HGPRP polypeptides.

Sequence 807 AA;

Query Match 78.0%; Score 32; DB 21; Length 807;
 Best Local Similarity 85.7%; Pred. No. 2.9e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
 I | I I I I I
 Db 205 mlptylk 211

RESULT 8

AAAY40440
 ID AAY40440 standard; Protein; 986 AA.

XX
 AC AAY40440;

XX 26-MAY-2000 (first entry)

XX Human brain-derived G-protein coupled receptor protein.

XX G-protein coupled receptor protein; human; brain; gene therapy;
 KW genetic disease; screening assay.

XX Homo sapiens.

XX WO200008053-A1.

XX 17-FEB-2000.

XX 05-AUG-1999; 99WO-JP04233.

XX 07-AUG-1998; 98JP-0225059.

XX (TAKE) TAKEDA CHEM IND LTD.

XX (KAZU-) KAZUSA DNA RES INST.

XX Ohara O, Nagase T, Nomura N, Mogi S, Yamamoto K, Kurokawa T;

XX WPI; 2000-195555/17.

XX N-PSDB; AAZ87930.

XX Human brain-derived G-protein coupled receptor protein and encoding DNA
 PT used in gene therapy, is also useful for e.g. identifying ligands and
 PT raising antibodies and antisera -

PS Claim 1; Fig 1; 80pp; Japanese.

XX This represents a human brain-derived G-protein coupled receptor
 CC protein. The polypeptide can be expressed by standard recombinant
 CC methodology. The novel G-protein coupled receptor protein can be used
 CC for identifying ligands; raising antibodies and antisera; developing

CC receptor-binding assay system; and screening for drug candidates. The
 CC encoding polynucleotide can be used in; probing for diagnostic genes;
 CC constructing PCR primers; making transgenic animals; and in gene therapy.

XX Sequence 986 AA;

Query Match 78.0%; Score 32; DB 21; Length 986;
 Best Local Similarity 85.7%; Pred. No. 3.6e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
 I | I I I I I
 Db 384 mlptylk 390

RESULT 9

AAW75457

ID AAW75457 standard; peptide; 8 AA.

XX AAW75457;

XX 27-APR-1999 (first entry)

XX Mammalian tub protein tyrosine phosphorylation site.

XX Human; wild type; tubby; identification; SH2 domain; mammal; obesity;
 KW body weight disorder; cachexia; anorexia.

XX Homo sapiens.

XX Mus sp.

XX US5861239-A.

XX 19-JAN-1999.

XX 02-SEP-1997; 97US-0922267.

XX 02-SEP-1997; 97US-0922267.

XX 12-APR-1996; 96US-0631200.

XX 28-MAR-1997; 97US-0829553.

XX (MILL-) MILLENNIUM PHARM INC.

XX Kapeller R, Kleyn PW, Moore KJ;

XX WPI; 1999-130383/11.

XX Identifying compounds which modulate tub protein activity - by
 PT detecting compounds which alter the interaction of tub protein with
 PT a SH2 containing peptide, used to develop agents for treating e.g.
 PT obesity, cachexia or anorexia

XX Disclosure; Column 13; 95pp; English.

XX This sequence represents a tyrosine phosphorylation site/SH2 docking
 CC domain found in the amino acid sequence of the mouse and human "tub"
 CC proteins (AAW75450 and AAW75451 respectively). The invention relates to
 CC a method for identifying compounds that modulate tub protein activity,
 CC especially its interaction with proteins containing an SH2 domain. The
 CC method can be used for identifying compounds which modulate tub protein
 CC activity for use in the treatment of mammalian body weight disorders
 CC including obesity, cachexia and anorexia.

XX Sequence 8 AA;

Query Match 75.6%; Score 31; DB 20; Length 8;
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTI 6
 I | I I I I I

Db 2 mypty 6

RESULT 10
AAW97073
ID AAW97073 standard; peptide; 12 AA.
AC AAW97073;
XX
XX 29-APR-1999 (first entry)
XX
XX Peptidomimetic capable of inhabiting CD28 and/or CTLA-4 interactions.
XX
XX Peptidomimetic; CD28; CTLA-4; CD80; CD86; B7-1; B7-2; diabetes;
KW immune system disease; autoimmune disease; psoriasis; multiple sclerosis;
KW lupus erythematosus; rheumatoid arthritis; transplant rejection; cancer.
XX Synthetic.
XX
XX WO9856401-A1.
XX
XX 17-DEC-1998.
XX
XX 11-JUN-1998; 98WO-US12312.
XX
XX 12-JUN-1997; 97US-0049470.
XX
XX (ISTF) ARS APPLIED RES SYSTEMS HOLDING NV.
XX
XX Blechner S, El Tayar N, Jameson B, Tepper M;
XX
XX WPI; 1999-080854/07.
XX
XX New peptidomimetic compounds - useful in the prophylaxis, diagnosis
PT and treatment of pathologies and disorders, which are improved by
PT inhibition of CD28 and/or CTLA-4 interaction with CD80 (B7-1) and
PT CD86 (B7-2)
XX
XX Disclosure; Page 9; 62pp; English.
XX
XX AAW97053-82 represent peptidomimetic compounds that inhibit CD28 and/or
CC CTLA-4 interactions with CD80 and CD86. The peptides are used in
CC pharmaceutical compositions for the prophylaxis, diagnosis and treatment
CC of pathologies and disorders, which are improved by inhibition of CD28
CC and/or CTLA-4 interaction with CD80 (B7-1) and CD86 (B7-2). Disorders
CC including immune system diseases (e.g. Autoimmune diseases such as
CC psoriasis, multiple sclerosis, lupus erythematosus, diabetes, rheumatoid
CC arthritis, and therapy against solid organ/cellular transplant
CC rejection) and cancer.

Sequence 12 AA;

Query Match 75.6%; Score 31; DB 20; Length 12;
Best Local Similarity 71.4%; Pred. No. 5.2;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 MYPTVLK 8
DB 2 myppylr 8

RESULT 11
AAV12222
ID AAV12222 standard; Protein; 58 AA.
AC AAV12222;
XX
XX 18-JUN-1999 (first entry)
XX
XX Human 5' EST secreted protein SEQ ID NO: 535.

KW Human; secreted protein; EST; expressed sequence tag; diagnosis;
KW forensic; gene therapy; chromosome mapping; signal peptide;
KW upstream regulatory sequence; cytokine activity; cell proliferation;
KW differentiation; haematopoiesis regulation; tissue growth regulation;
KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; anti-inflammatory; tumour inhibition.
XX
XX Homo sapiens.
XX
XX WO9906554-A2.
XX
XX 11-FEB-1999.
XX
XX 31-JUL-1998; 98WO-IB01238.
XX
XX 01-AUG-1997; 97US-0905134.
XX
XX (GEST) GENSET.
XX
XX Duclert A, Dumas Milne Edwards J, Lacroix B;
XX
XX WPI; 1999-153784/13.
XX
XX N-PSDB; AAX41055.
XX
XX New nucleic acids encoding human secreted proteins - obtained from
PT cDNA libraries prepared from kidney, fetal kidney, dystrophic
PT muscle, muscle and heart tissue
XX
XX Claim 34; Page 598; 622pp; English.

XX
XX AAX40826 to AAX41093 represent 5' expressed sequence tags (ESTs) for
CC human secreted proteins and encode the proteins given in AAY01602 and
CC AAY11994 to AAY12260, respectively. The proteins given represent the
CC signal peptide and an N-terminal fragment of a secreted protein. The
CC nucleic acid sequences can be used for producing secreted human gene
CC products. They can also be used to develop products for diagnosis and
CC therapy. The proteins obtained may have cytokine activity, cell
CC proliferation/differentiation activity, haematopoiesis regulating
CC activity, tissue growth regulating activity, reproductive hormone
CC regulating activity, chemotactic/chemokinetic activity, haemostatic and
CC thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, tumour inhibition activity or other activities. The products
CC can be used in forensic, gene therapy and chromosome mapping procedures.
CC The sequences can also be used for obtaining coding sequences.
CC sequences. The nucleic acids encoding the signal peptide can be used
CC for directing extracellular secretion of a polypeptide or the insertion
CC of a polypeptide into a membrane, or importing a polypeptide into
CC a cell.

XX
XX Sequence 58 AA;

Query Match 75.6%; Score 31; DB 20; Length 58;
Best Local Similarity 83.3%; Pred. No. 28;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYL 7
DB 1 mypsyl 6

RESULT 12
AAB74618
ID AAB74618 standard; protein; 173 AA.
XX
XX AAB74618;
XX
XX 22-MAY-2001 (first entry)
XX
XX Human 72 kDa gelatinase catalytic domain protein sequence SEQ ID NO:8.
XX
XX Human; 92 kDa gelatinase; 72 kDa gelatinase; matrilysin; gelatinase;
KW catalytic domain; 92 CD; 72 CD; matrilysin CD; elastin; vulnery;

KW excess connective tissue removal; dermatological; keloid; scleroderma;
 KW post-operative fibrosis; intervertebral disc injection; fibrotic disease;
 KW hypertrophic scar; wound debridement; post-surgical adhesion;
 KW idiopathic pulmonary fibrosis.

OS Homo sapiens.

XX US6194189-B1.

PN 27-FEB-2001.

XX 16-DEC-1994; 94US-0357820.

XX 16-DEC-1994; 94US-0357820.

XX (UNIW) UNIV WASHINGTON.

PA Senior RM;

PI WPI; 2001-243407/25.

XX New gelatinase truncated mutant useful for treating disorders requiring
 CC the removal of excess connective tissue, e.g. keloids, post-operative
 PT fibrosis, intervertebral disc injections, hypertrophic scars -
 XX Example; Column 13-16; l1pp; English.

PS The present invention describes a cDNA sequence which encodes the

CC truncated mutant of the 92 kDa gelatinase having an amino acid sequence
 CC as given in AAB74617 consisting of residues 106-216 fused to residues
 CC 391-443 of the parent molecule. The truncated mutant has dermatological
 CC and vulnary activities. The truncated mutant is useful for treating
 CC disorders requiring the removal of excess connective tissue,
 CC e.g., keloids, post-operative fibrosis, intervertebral disc injections,
 CC hypertrophic scars, wound debridement, post-surgical adhesions and
 CC various fibrotic diseases (including scleroderma, idiopathic pulmonary
 CC fibrosis). The truncated mutant is catalytically active compared to
 CC the full protein. Unlike the full protein, the truncated mutant is
 CC essentially inactive against insoluble elastin, and does not require
 CC activation to be enzymatically active. The present sequence represents
 CC a 72 kDa gelatinase catalytic domain (72 CD) containing protein, which
 CC is given in the exemplification of the present invention.

XX Sequence 173 AA;

Query Match 75.6%; Score 31; DB 22; Length 173;

Best Local Similarity 100.0%; Pred. No. 89; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6

IIIII

Db 138 mypty 142

RESULT 13

AAB68682

ID AAB68682 standard; Protein; 173 AA.

XX AAB68682;

XX 02-MAY-2001 (first entry)

XX Matrilysin catalytic domain.

XX Gelatinase; excess connective tissue removal; matrilysin;

KW matrix metalloproteinase.

XX Unidentified.

OS US6184021-B1.

PN 06-FEB-2001.

XX

XX 19-MAY-1995; 95US-0444528.

XX 16-DEC-1994; 94US-0357820.

XX (UNIW) UNIV WASHINGTON.

XX Senior RM;

XX WPI; 2001-202001/20.

XX New truncated mutant of 92 kDa gelatinase which is catalytically
 CC active, but is inactive against insoluble elastin, useful for treating
 PT disorders requiring the removal of excess connective tissues such as
 PT keloids -

XX Examples: Fig 2; l1pp; English.

XX The present invention relates to a truncated mutant (92 CD) of the 92
 CC kDa gelatinase (see AAB68681). The truncated protein is
 CC useful for treating disorders requiring the removal of excess connective
 CC tissue, e.g. keloids, post-operative fibrosis, intervertebral disc
 CC injections, hypertrophic scars, wound debridement, post-surgical
 CC adhesions and various fibrotic diseases (scleroderma, idiopathic
 CC pulmonary fibrosis). Gelatinase is a matrix metalloproteinase and is also
 CC known as gelatinase B and MMP-9. The truncated protein is catalytically
 CC active comparable to the full protein but unlike the full protein is
 CC essentially inactive against insoluble elastin. The present sequence is
 CC the catalytic domain of matrilysin, which was used in a sequence homology
 CC alignment with the 92 CD protein of the present invention. Matrilysin is
 CC also a matrix metalloproteinase.

XX Sequence 173 AA;

Query Match 75.6%; Score 31; DB 22; Length 173;

Best Local Similarity 100.0%; Pred. No. 89;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6

IIIII

Db 138 mypty 142

RESULT 14

AAW76253

ID AAW76253 standard; Protein; 261 AA.

XX AAW76253;

XX 02-DEC-1998 (first entry)

XX Human matrilysin-like protein.

XX Matrilysin; inhibitor; tumour necrosis factor; screening; MMP; Infection;
 KW matrix metalloproteinase; antagonist; diagnosis; underexpression; cancer;
 KW degeneration; extracellular matrix; arthritis; cardiovascular disease;
 KW cachexia; multiple sclerosis.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..22

FT Protein /label= signal

FT /label= 23..261

FT /label= matrilysin_like_protein

XX WO9831818-A2.

XX 23-JUL-1998.

XX 20-JAN-1998; 98WO-0500783.

XX

PR 01-AUG-1997; 97US-0054541.
 PR 21-JAN-1997; 97US-0034205.
 PR 13-JUN-1997; 97US-0049607.
 PA (HUMA-) HUMAN GENOME SCI INC.

PI Brewer L, Gentz R, Ni J, Rosen CA, Ruben SM;
 DR WPI: 1998-414114/35.
 DR N-PSDB; AAV61633.

XX Isolated nucleic acid encoding human metallo-protease(s) - used for
 PT diagnosis, treatment and prevention of, e.g. cancer, inflammation,
 PT neurological disease and infections
 PS Claim 20a; Fig 4; 81pp; English.

XX This sequence represents a novel human matrilysin-like protein which is
 CC an inhibitor of the members of the matrix metalloproteinase, MMP,
 CC family. This protein can be used in assays to screen for agonists and
 CC antagonists and the nucleic acid is used as a probe for gene mapping, in
 CC situ hybridisation and detection of corresponding genes in human tissue,
 CC and as sources of probes and primers for diagnosis. The protein and its
 CC antigenic fragments are used to raise antibodies (Ab) (which can be used
 CC for diagnosis in usual immunoassays or for in vivo imaging) and to screen
 CC for (ant)agonists. Agonists can be used to treat cancer or other
 CC conditions associated with underexpression of TNF-alpha. Antagonists are
 CC used to treat diseases associated with degeneration of the extracellular
 CC matrix (cancer, arthritis, cardiovascular disease, cachexia and multiple
 CC sclerosis).

XX Sequence 261 AA;

Query Match 75.6%; Score 31; DB 19; Length 261;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
 DB 226 mypty 230

RESULT 15

AAW76254
 ID AAW76254 standard; Protein; 264 AA.

XX AAW76254;

XX 02-DEC-1998 (first entry)
 DE Human matrilysin protein.

XX Matrilysin; inhibitor; tumour necrosis factor; screening; MMP; infection;
 KW matrix metalloproteinase; antagonist; diagnosis; underexpression; cancer;
 KW degeneration; extracellular matrix; arthritis; cardiovascular disease;
 XX cachexia; multiple sclerosis.

OS Homo sapiens.

XX WO9831818-A2.

PN 23-JUL-1998.

XX 20-JAN-1998; 98WO-US00783.

XX 01-AUG-1997; 97US-0054541.

PR 21-JAN-1997; 97US-0034205.

XX 13-JUN-1997; 97US-0049607.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Brewer L, Gentz R, Ni J, Rosen CA, Ruben SM;

XX WPI: 1998-414114/35.
 DR Isolated nucleic acid encoding human metallo-protease(s) - used for
 PT diagnosis, treatment and prevention of, e.g. cancer, inflammation,
 PT neurological disease and infections
 XX Disclosure; Fig 5; 81pp; English.

XX This sequence represents the human matrilysin protein which is an
 CC inhibitor of the members of the matrix metalloproteinase, MMP, family.
 CC This protein is used in the identification of a novel human
 CC matrilysin-like protein which can be used in assays to screen for
 CC agonists and antagonists and the nucleic acid is used as a probe for gene
 CC mapping, in situ hybridisation and detection of corresponding genes in
 CC human tissue and as sources of probes and primers for diagnosis. The
 CC protein and its antigenic fragments are used to raise antibodies (Ab)
 CC (which can be used for (ant)agonists. Agonists can be used to treat
 CC cancer or other conditions associated with underexpression of TNF-alpha.
 CC Antagonists are used to treat diseases associated with degeneration of
 CC the extracellular matrix (cancer, arthritis, cardiovascular disease,
 CC cachexia and multiple sclerosis).

XX Sequence 264 AA;

Query Match 75.6%; Score 31; DB 19; Length 264;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
 DB 232 mypty 236

Search completed: September 4, 2001, 15:50:26
 Job time: 164 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2001, 15:51:13 ; Search time 40.98 Seconds
(without alignments)
4.020 Million cell updates/sec

Title: US-09-630-345-6
Perfect score: 41 XMYPTYLK 8
Sequence: 1

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	97.6	8	1	US-07-620-410-3
2	31	75.6	8	2	US-08-922-267A-65
3	31	75.6	173	4	US-08-444-628-8
4	31	75.6	173	3	US-08-357-820-8
5	31	75.6	261	3	US-09-009-156-5
6	31	75.6	264	3	US-09-009-156-6
7	31	75.6	267	4	US-08-448-489-18
8	31	75.6	271	3	US-08-896-062-2
9	31	75.6	459	1	US-08-630-592-2
10	31	75.6	459	1	US-08-714-991-2
11	31	75.6	459	3	US-09-032-365A-2
12	31	75.6	460	1	US-08-630-592-7
13	31	75.6	460	1	US-08-714-991-7
14	31	75.6	460	3	US-09-032-365A-8
15	31	75.6	482	4	US-08-068-392-3
16	31	75.6	482	4	US-08-396-988-3
17	31	75.6	505	1	US-08-631-200-2
18	31	75.6	505	1	US-08-630-592-4
19	31	75.6	505	1	US-08-714-991-4
20	31	75.6	505	2	US-08-829-553-2
21	31	75.6	505	2	US-08-922-267A-2
22	31	75.6	505	2	US-08-936-707A-2
23	31	75.6	505	2	US-08-936-706A-2
24	31	75.6	505	3	US-09-248-203-2
25	31	75.6	505	3	US-09-032-365A-4
26	31	75.6	505	4	US-08-812-824-3
27	31	75.6	505	4	US-09-406-071-2

QY 2 MYPTYLK 8
|||||||

Query Match 97.6%; Score 40; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ALIGNMENTS

RESULT 1
US-07-620-410-3
: Sequence 3, Application US/07620410
: Patent No. 5217953
: GENERAL INFORMATION:
: APPLICANT: Gozes, Iillana
: APPLICANT: Brennenman, Douglas E.
: APPLICANT: Fridkin, Matl
: APPLICANT: Moody, Terry
: TITLE OF INVENTION: A NOVEL VASOACTIVE INTESTINAL PEPTIDE
: TITLE OF INVENTION: ANTAGONIST
: NUMBER OF SEQUENCES: 3
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
: STREET: 1615 L. Street, N.W.
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20036
: COMPUTER READABLE FORM: disk
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/620,410
: FILING DATE: 19901130
: CLASSIFICATION: 514
: ATTORNEY/AGENT INFORMATION:
: NAME: Scott, Watson T.
: REGISTRATION NUMBER: 26,581
: REFERENCE/DOCKET NUMBER: WTS/5683/82679
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202) 861-3000
: TELEFAX: (202) 822-0944
: TELEX: 248453 CUSH
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 8 amino acids
: TYPE: AMINO ACID
: TOPOLOGY: linear
US-07-620-410-3

28	31	75.6	506	1	US-08-631-200-8	Sequence 8, Appli
29	31	75.6	506	2	US-08-829-553-8	Sequence 8, Appli
30	31	75.6	506	2	US-08-922-267A-8	Sequence 8, Appli
31	31	75.6	506	2	US-08-936-707A-8	Sequence 8, Appli
32	31	75.6	506	2	US-08-936-706A-8	Sequence 8, Appli
33	31	75.6	506	3	US-09-248-203-8	Sequence 8, Appli
34	31	75.6	506	3	US-09-032-365A-62	Sequence 62, Appli
35	31	75.6	506	4	US-08-812-824-4	Sequence 4, Appli
36	31	75.6	506	4	US-09-406-071-8	Sequence 8, Appli
37	31	75.6	512	3	US-09-032-365A-60	Sequence 60, Appli
38	31	75.6	518	3	US-09-032-365A-58	Sequence 58, Appli
39	31	75.6	561	1	US-08-714-991-27	Sequence 27, Appli
40	31	75.6	561	3	US-09-032-365A-10	Sequence 10, Appli
41	30	73.2	73	1	US-08-379-538-7	Sequence 7, Appli
42	30	73.2	452	2	US-08-731-079A-2	Sequence 2, Appli
43	30	73.2	547	4	US-09-178-252-11	Sequence 11, Appli
44	30	73.2	725	1	US-08-448-170-4	Sequence 4, Appli
45	30	73.2	725	4	US-08-961-803-7	Sequence 7, Appli

Db 2 MYPTYLK 8

RESULT 2

US-08-922-267A-65
; Sequence 65, Application US/08922267A
; Patent No. 5861239
; GENERAL INFORMATION:
; APPLICANT: Klevy, Patrick W.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; NUMBER OF SEQUENCES: 82
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/922,267A
FILING DATE: 2-SEP-1997
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/829,553
FILING DATE: 28-MAR-1997
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/631,200
FILING DATE: 12-APR-1996
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-085
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 8 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: peptide
MOLECULE TYPE: peptide
US-08-922-267A-65

Query Match 75.6%; Score 31; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6

Db 2 MYPTY 6

RESULT 3

US-08-444-628-8
; Sequence 8, Application US/08444628
; Patent No. 6184021
; GENERAL INFORMATION:
; APPLICANT: Senior, Robert M.
; TITLE OF INVENTION: Catalytically-active Gelatinase Mutant
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Scott J. Meyer, Monsanto/Searle, A3SG
STREET: 800 N. Lindbergh Blvd.
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63167

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/444,628
FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/357,820
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Meyer, Scott J.
REGISTRATION NUMBER: 25,275
REFERENCE/DOCKET NUMBER: WU-2855
TELEPHONE: (314)694-3117
TELEFAX: (314)694-5435
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 173 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-444-628-8

Query Match 75.6%; Score 31; DB 4; Length 173;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6

Db 138 MYPTY 142

RESULT 4

US-08-357-820-8
; Sequence 8, Application US/08357820
; Patent No. 6194189
; GENERAL INFORMATION:
; APPLICANT: Senior, Robert M.
; TITLE OF INVENTION: Catalytically-active Gelatinase Mutant
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Scott J. Meyer, Monsanto/Searle, A3SG
STREET: 800 N. Lindbergh Blvd.
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63167

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/357,820
FILING DATE:

CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Meyer, Scott J.
REGISTRATION NUMBER: 25,275
REFERENCE/DOCKET NUMBER: WU-2855
TELECOMMUNICATION INFORMATION:

TELEPHONE: (314)694-3117
TELEFAX: (314)694-5435
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 173 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-357-820-8

Query Match 75.6%; Score 31; DB 4; Length 173;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
Db 138 MYPTY 142

RESULT 5

US-09-009-156-5
Sequence 5, Application US/09009156
Patent No. 6046031
GENERAL INFORMATION:
APPLICANT: NI, Jian
APPLICANT: Ruben, Steven
APPLICANT: Brewer, Laurie
APPLICANT: Gentz, Reiner
TITLE OF INVENTION: No. 6046031le1 Metalloproteinases
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSER: Human Genome Sciences, Inc
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: US
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/009,156
FILING DATE:

CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/034,205
FILING DATE: 21-JAN-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/049,607
FILING DATE: 13-JUN-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/054,541
FILING DATE: 01-AUG-1997
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PF376
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-309-8504
TELEFAX: 301-309-8439
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 261 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-009-156-5

Query Match 75.6%; Score 31; DB 3; Length 261;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
Db 226 MYPTY 230

RESULT 6

US-09-009-156-6
Sequence 6, Application US/09009156
Patent No. 6046031
GENERAL INFORMATION:
APPLICANT: NI, Jian
APPLICANT: Ruben, Steven
APPLICANT: Brewer, Laurie
APPLICANT: Gentz, Reiner
TITLE OF INVENTION: No. 6046031le1 Metalloproteinases
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSER: Human Genome Sciences, Inc
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: US
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/009,156
FILING DATE:

CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/034,205
FILING DATE: 21-JAN-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/049,607
FILING DATE: 13-JUN-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/054,541
FILING DATE: 01-AUG-1997
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PF376
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-309-8504
TELEFAX: 301-309-8439
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 264 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-009-156-6

Query Match 75.6%; Score 31; DB 3; Length 264;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
Db 232 MYPTY 236

RESULT 7

US-08-448-489-18
; Sequence 18, Application US/08448489
; Patent No. 6184022
; GENERAL INFORMATION:
; APPLICANT: SEIKI, Motoharu
; APPLICANT: SATO, Hitoshi
; APPLICANT: SHINAGAWA, Akira
; TITLE OF INVENTION: NOVEL METALLOPROTEINASE AND ENCODING DNA THEREFOR
; FILE REFERENCE: 55-290P
; CURRENT APPLICATION NUMBER: US/08/448,489
; CURRENT FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 18
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Known Member of
; OTHER INFORMATION: Matrix Metalloproteinase Family
; US-448-489-18

Query Match 75.6%; Score 31; DB 4; Length 267;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
Db 232 MYPTY 236

RESULT 8
US-08-896-062-2
; Sequence 2, Application US/08896062
; Patent No. 6010893
; GENERAL INFORMATION:
; APPLICANT: KIHARA, Yasunori
; TITLE OF INVENTION: PROCESS FOR PRODUCING HUMAN MATRILYSIN
; TITLE OF INVENTION: BY MEANS OF RECOMBINANT DNA
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Broadway and Neimark
; STREET: 419 Seventh Street N.W., Ste. 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States of America
; ZIP: 20004

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/896,062
FILING DATE: 17-JUL-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/530,984
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Broadway, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: KIHARA-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
TELEX: 248633

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 271 amino acids
TYPE: amino acid
TOPOLOGY: linear

US-08-896-062-2
; MOLECULE TYPE: protein
; US-08-896-062-2

Query Match 75.6%; Score 31; DB 3; Length 271;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
Db 236 MYPTY 240

RESULT 9
US-08-630-592-2
; Sequence 2, Application US/08630592
; Patent No. 5770432
; GENERAL INFORMATION:
; APPLICANT: Nishina, Patsy
; APPLICANT: No. 5770432entrauth, Konrad
; APPLICANT: Nagert, Juergen
; APPLICANT: No. 5770432th, Michael
; TITLE OF INVENTION: Obesity Associated Genes
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
; STREET: 3400 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 941114187

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MSDOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/630,592
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sherwood, Pamela J.
REGISTRATION NUMBER: 36,677
REFERENCE/DOCKET NUMBER: A59504/BIR/PJS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 7811989
TELEFAX: (415) 3983249
TELEX: 910 277299

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 459 amino acids
TYPE: amino acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-630-592-2

Query Match 75.6%; Score 31; DB 1; Length 459;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
Db 233 MYPTY 237

RESULT 10
US-08-714-991-2
; Sequence 2, Application US/08714991
; Patent No. 5776762
; GENERAL INFORMATION:
; APPLICANT: NORTH, Michael

APPLICANT: NISHINA, Patsy
APPLICANT: NO. 5776762en-Trauth, Konrad
APPLICANT: NAGGERT, Juergen
TITLE OF INVENTION: OBESITY ASSOCIATED GENES
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/714,991
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: SHERWOOD, Pamela J.
REGISTRATION NUMBER: 36,677
REFERENCE/DOCKET NUMBER: A-59504-1/PJS
TELEPHONE: 415-494-8700
TELEFAX: 415-494-8771
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 459 amino acids
TYPE: amino acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-714-991-2

Query Match 75.6%; Score 31; DB 1; Length 459;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
|||||
Db 233 MYPTY 237

ULT 11
09-032-365A-2
Sequence 2, Application US/09032365A
Patent No. 6114502
GENERAL INFORMATION:
APPLICANT: NO. 6114502th, Michael
APPLICANT: Nishina, Patsy
APPLICANT: NAGGERT, Juergen
APPLICANT: NO. 6114502en-Trauth, Konrad
TITLE OF INVENTION: GENE FAMILY ASSOCIATED WITH
TITLE OF INVENTION: NEUROSENSORY DEFECTS
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bozicevic & Reed, LLP
STREET: 285 Hamilton Avenue, Suite 200
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/032,365A
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Sherwood, Pamela J.
REGISTRATION NUMBER: 36,677
REFERENCE/DOCKET NUMBER: SEQ-2CIP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-327-3400
TELEFAX: 650 327-3231
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 459 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-032-365A-2

Query Match 75.6%; Score 31; DB 3; Length 459;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
|||||
Db 233 MYPTY 237

RESULT 12
US-08-630-592-7
Sequence 7, Application US/08630592
Patent No. 5770432
GENERAL INFORMATION:
APPLICANT: Nishina, Patsy
APPLICANT: NO. 5770432enTrauth, Konrad
APPLICANT: NAGGERT, Juergen
APPLICANT: NO. 5770432th, Michael
TITLE OF INVENTION: Obesity Associated Genes
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
STREET: 3400 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 941114187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PCDOS/MSDOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/630,592
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sherwood, Pamela J.
REGISTRATION NUMBER: 36,677
REFERENCE/DOCKET NUMBER: A59504/BIR/PJS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 7811989
TELEFAX: (415) 3983249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 460 amino acids
TYPE: amino acid
STRANDEDNESS: double

TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-630-592-7

Query Match 75.6% Score 31; DB 1; Length 460;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 2 MYPTY 6
Db 234 MYPTY 238

RESULT 13
US-08-714-991-7

; Sequence 7, Application US/08714991
; Patent No. 5776762

; GENERAL INFORMATION:

APPLICANT: NISHINA, Michael

APPLICANT: NAGGERT, Juergen

APPLICANT: No. 5776762en-Trauth, Konrad

TITLE OF INVENTION: OBESITY ASSOCIATED GENES

NUMBER OF SEQUENCES: 28

CORRESPONDENCE ADDRESS:

ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT

STREET: 4 Embarcadero Center, Suite 3400

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-4187

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/714,991

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: SHERWOOD, Pamela J.

REGISTRATION NUMBER: 36,677

REFERENCE/DOCKET NUMBER: A-59504-1/PJS

TELEPHONE: 415-494-8700

TELEFAX: 415-494-8771

TELEX: 910 277299

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 460 amino acids

TYPE: amino acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-714-991-7

Query Match 75.6% Score 31; DB 1; Length 460;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 2 MYPTY 6
Db 234 MYPTY 238

RESULT 14

US-09-032-365A-8

; Sequence 8, Application US/09032365A

; Patent No. 6114502

; GENERAL INFORMATION:

APPLICANT: No. 6114502th, Michael
APPLICANT: Nishina, Patsy
APPLICANT: Naggart, Juergen
APPLICANT: No. 6114502en-Trauth, Konrad
TITLE OF INVENTION: GENE FAMILY ASSOCIATED WITH
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:

ADDRESSEE: Bozicevic & Reed, LLP

STREET: 285 Hamilton Avenue, Suite 200

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94301

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/032,365A

FILING DATE:

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Sherwood, Pamela J

REGISTRATION NUMBER: 36,677

REFERENCE/DOCKET NUMBER: SEQ-2CIP2

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-327-3400

TELEFAX: 650 327-3231

TELEX:

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 460 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-032-365A-8

Query Match 75.6% Score 31; DB 3; Length 460;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 2 MYPTY 6
Db 234 MYPTY 238

RESULT 15

US-08-068-392-3

; Sequence 3, Application US/08068392

; Patent No. 6150152

; GENERAL INFORMATION:

APPLICANT: Shapiro, Steven M.

TITLE OF INVENTION: Human Macrophage Metalloproteinase

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:

ADDRESSEE: Scott J. Meyer, Monsanto Co., A3SM

STREET: 800 N. Lindbergh Blvd.

CITY: St. Louis

STATE: Missouri

COUNTRY: USA

ZIP: 63167

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/068,392
;; FILING DATE: 19930528
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Meyer, Scott J.
;; REGISTRATION NUMBER: 25275
;; REFERENCE/DOCKET NUMBER: 07-24(12406)A
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (314)694-3117
;; TELEFAX: (314)694-5435
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 462 amino acids
;; TYPE: AMINO ACID
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-08-068-392-3

Query Match 75.6%; Score 31; DB 4; Length 462;
Best Local Similarity 100.0%; Pred. No. 1.12e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPT 6
|
|
|
|
|
Db 229 MYPT 233

Search completed: September 4, 2001, 15:51:13
Job time: 196 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2001, 15:52:09 ; Search time 49.08 Seconds
(without alignments)
12.416 Million cell updates/sec

Title: US-09-630-345-6
Perfect score: 41
Sequence: 1 MYPTYLK 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37	90.2	384	I51267	collagenase (EC 3.4.24.24) - bullfrog
2	35	85.4	393	B86644	transporter ybfb [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
3	34	82.9	436	C69764	4-aminobutyrate am
4	33	80.5	218	E86750	hypothetical prote
5	33	80.5	269	D69011	hypothetical prote
6	32	78.0	340	MMBE5	cell fusion protei
7	32	78.0	349	JU0458	peroxidase (EC 1.1
8	32	78.0	593	JC4884	organic cation tra
9	32	78.0	1263	F44466	DNA-directed RNA p
10	31	75.6	247	B85515	hypothetical prote
11	31	75.6	262	T29698	hypothetical prote
12	31	75.6	267	KCHUM	matrilysin (EC 3.4
13	31	75.6	267	A57490	matrilysin (EC 3.4
14	31	75.6	301	S57923	SEC14 protein - ye
15	31	75.6	378	F86307	hypothetical prote
16	31	75.6	385	E86359	Similar to seed ma
17	31	75.6	462	A42401	macrophage elastas
18	31	75.6	478	A86677	phospho-beta-gluc
19	31	75.6	483	JC5743	matrix metallopro
20	31	75.6	495	PLWLB	L1 protein - bovin
21	31	75.6	501	PLWLB2	L1 protein - bovin
22	31	75.6	501	PLWLB2	L1 protein - bovin
23	31	75.6	502	T19708	hypothetical prote
24	31	75.6	503	T29175	hypothetical prote
25	31	75.6	505	S68518	tub protein, brain
26	31	75.6	511	T19496	hypothetical prote
27	31	75.6	513	PLWLDP	L1 protein - deer
28	31	75.6	516	T15633	hypothetical prote
29	31	75.6	565	H69363	hypothetical prote

30	31	75.6	662	2	S42826	probable ATPase -
31	31	75.6	2185	1	GNNYSV	genome polyprotein
32	31	75.6	2185	1	GNNYSV	genome polyprotein
33	30	73.2	141	2	T29507	hypothetical prote
34	30	73.2	152	2	G83476	hypothetical prote
35	30	73.2	155	2	T17986	hypothetical prote
36	30	73.2	192	2	T11930	NADH dehydrogenase
37	30	73.2	239	2	B64757	glycolate oxidase
38	30	73.2	239	2	D85522	probable dehydroge
39	30	73.2	248	2	S57910	probable response
40	30	73.2	262	2	C81384	shikimate 5-dehydr
41	30	73.2	287	2	D83414	probable transcrip
42	30	73.2	343	1	MMBEA5	cell fusion protei
43	30	73.2	368	1	HLBECM	membrane glycoprot
44	30	73.2	410	2	T20397	hypothetical prote
45	30	73.2	491	2	JE0396	phospho-beta-galac

ALIGNMENTS

RESULT 1

I51267
collagenase (EC 3.4.24.24) - bullfrog
C:Species: Rana catesbeiana (bullfrog)
C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 22-Jun-1999
C:Accession: I51267
R:Oofusa, K.; Yomori, S.; Yoshizato, K.
Int. J. Dev. Biol. 38, 345-350, 1994
A>Title: Regionally and hormonally regulated expression of genes of collagen and coll
A:Reference number: I51267; MUID:95071832
A:Accession: I51267
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-384 <OOF>
A:Cross-references: GB:S75623; NID:9913070; PIDN:AAB32661.1; PID:9913071
C:Speciesfamily: interstitial collagenase; hemopexin repeat homology; matrix metallopro
C:Keywords: hydrolase; metalloproteinase; zinc; zymogen
F:53-231/Domain: matrix metalloproteinase homology <MMP>
F:236-381/Domain: hemopexin repeat homology <PNX>
F:81,189,193,199/Binding site: zinc, catalytic (Cys, His, His, His) (inhibited) #stat
F:189,193,199/Binding site: zinc, catalytic (His) (active) #status predicted
F:190/Active site: Glu #status predicted

Query Match 90.2%; Score 37; DB 2; Length 384;
Best Local Similarity 85.7%; Pred. No. 5;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
|||||
DB 207 MYPTVLR 213

RESULT 2

B86644
transporter ybfb [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C:Species: Lactococcus lactis subsp. lactis
C>Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 23-Mar-2001
C:Accession: B86644
R:Botolin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissensbach, J.; Eh
Genome Res. in press, 2001
A>Title: The complete genome sequence of the lactic acid bacterium.
A:Reference number: A86625
A:Accession: B86644
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-393 <STO>
A:Cross-references: GB:AE005176; NID:g12723004; PIDN:AAK04252.1; GSPDB:GN00146
C:Genetics:
A:Gene: ybfb

Query Match 85.4%; Score 35; DB 2; Length 393;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 MYPTYL 7
 |||||
 Db 278 MYPTYL 283

RESULT 3
 C69764
 A:amino butyrate aminotransferase homolog ycnG - Bacillus subtilis
 C:Species: Bacillus subtilis
 C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
 C:Accession: C69764
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertea, C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chao, A.; Ehrlich, S.D.; Emmer, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.; Figue 390, 249-256, 1997
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallen, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F.; Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon, A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron, akenchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K. A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
 A:Reference number: A69580; MUID:98044033
 A:Accession: C69764
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-436 <KUN>
 A:Cross-references: GB:299106; GB:AL009126; NID:g2632653; PIDN:CAB12198.1; PID:g2632691
 A:Experimental source: strain 168
 C:Genetics:
 A:Gene: ycnG
 C:Superfamily: ornithine--oxo-acid aminotransferase

Query Match 82.9%; Score 34; DB 2; Length 436;
 Best Local Similarity 71.4%; Pred. No. 23;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 MYPTYL 8
 |||||
 88 MYPTYL 94

RESULT 4
 E86750
 hypothetical protein ykdb [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
 C:Species: Lactococcus lactis subsp. lactis
 C>Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 23-Mar-2001
 C:Accession: E86750
 R:Boletun, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarre, K.; Weissenbach, J.; Ehrlich Genome Res. In press, 2001
 A:Title: The complete genome sequence of the lactic acid bacterium.
 A:Reference number: A86625
 A:Accession: E86750
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-218 <STO>
 A:Cross-references: GB:AE005176; NID:g12723952; PIDN:AAK05103.1; GSPDB:GN00146
 A:Experimental source: strain IL1403
 C:Genetics:
 A:Gene: ykdb

Query Match 80.5%; Score 33; DB 2; Length 218;

Best Local Similarity 83.3%; Pred. No. 19;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 YPTYLK 8
 |||||
 Db 204 YPTYLK 209

RESULT 5
 D69011
 hypothetical protein MTH1087 - Methanobacterium thermoautotrophicum (strain Delta H)
 C:Species: Methanobacterium thermoautotrophicum
 C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
 C:Accession: D69011
 R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwan, K.; S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N. J. Bacteriol. 179, 7135-7155, 1997
 A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: fu
 A:Reference number: A69000; MUID:98037514
 A:Accession: D69011
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-269 <MTH>
 A:Cross-references: GB:AE000879; GB:AE000666; NID:g2622175; PIDN:AAB85576.1; PID:g262
 A:Experimental source: strain Delta H
 C:Genetics:
 A:Gene: MTH1087

Query Match 80.5%; Score 33; DB 2; Length 269;
 Best Local Similarity 83.3%; Pred. No. 23;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 YPTYLK 8
 |||||
 Db 175 YPTYLK 180

RESULT 6
 MMBE5
 cell fusion protein precursor - human herpesvirus 3
 C:Species: human herpesvirus 3, varicella-zoster virus
 C>Date: 30-Sep-1988 #sequence_revision 30-Sep-1988 #text_change 16-Jul-1999
 C:Accession: E27212
 R:Davidson, A.J.; Scott, J.E.
 J. Gen. Virol. 67, 1759-1816, 1986
 A:Title: The complete DNA sequence of varicella-zoster virus.
 A:Reference number: A27345; MUID:86306657
 A:Accession: E27212
 A:Molecule type: DNA
 A:Residues: 1-340 <DAV>
 A:Cross-references: EMBL:X04370; NID:g59989; PIDN:CAA27888.1; PID:g59994
 C:Genetics:
 A:Gene: 5
 C:Superfamily: herpesvirus cell fusion protein
 C:Keywords: membrane fusion; transmembrane protein
 F:1-31/Domain: signal sequence #status predicted <SIG>
 F:32-340/Product: cell fusion protein #status predicted <CFP>
 F:115-137/Domain: transmembrane #status predicted <TM1>
 F:220-238/Domain: transmembrane #status predicted <TM2>
 F:251-269/Domain: transmembrane #status predicted <TM3>
 F:307-322/Domain: transmembrane #status predicted <TM5>

Query Match 78.0%; Score 32; DB 1; Length 340;
 Best Local Similarity 57.1%; Pred. No. 47;
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 MYPTYLK 8
 |||||
 Db 244 LYPTYL 250

RESULT 7

JU0458
peroxidase (EC 1.11.1.7) E - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 04-Mar-2000
C:Accession: JU0458
R:Intaprak, C.; Higashimura, N.; Yamamoto, K.; Okada, N.; Shimmyo, A.; Takano, M.
Gene 98, 237-241, 1991
A:Title: Nucleotide sequences of two genomic DNAs encoding peroxidase of Arabidopsis thaliana
A:Reference number: JU0457; MUID:91200671
A:Accession: JU0458
A:Molecule type: DNA
A:Residues: 1-349 <INT>
A:Cross-references: GB:M58381; NID:gl166806; PIDN:AAA32842.1; PID:gl166807
C:Genetics:
A:Gene: prxEa
A:Introns: 76/3; 140/3; 197/1
C:Superfamily: peroxidase
Keywords: chromoprotein; heme; iron; metalloprotein; oxidoreductase
F:10-120/Disulfide bonds: #status predicted
F:67/Active site: Arg #status predicted
F:71,199/Binding site: heme iron (His) (axial ligands) #status predicted
F:73-78/Disulfide bonds: #status predicted
F:126-329/Disulfide bonds: #status predicted
F:206-238/Disulfide bonds: #status predicted

Query Match 78.0%; Score 32; DB 2; Length 349;

Best Local Similarity 83.3%; Pred. No. 48;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYL 7

Db 226 LYPTYL 231
:|||||

RESULT 8

JC4884
organic cation transporter protein 2 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 10-Sep-1996 #sequence_revision 18-Oct-1996 #text_change 05-Nov-1999
C:Accession: JC4884
R:Okuda, M.; Saito, H.; Urakami, Y.; Takano, M.; Inui, K.
Biochem. Biophys. Res. Commun. 224, 500-507, 1996
A:Title: cDNA cloning and functional expression of a novel rat kidney organic cation transporter
A:Reference number: JC4884; MUID:96295517
A:Accession: JC4884

A:Molecule type: mRNA

Residues: 1-593 <OKU>
A:Cross-references: DDBJ:D83044; NID:gl502282; PIDN:BAAL1754.1; PID:dl012421; PID:gl5022
A:Experimental source: kidney
C:Comment: This protein is responsible for the transport of cationic drugs in kidney.

Query Match 78.0%; Score 32; DB 2; Length 593;

Best Local Similarity 57.1%; Pred. No. 82;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYLK 8

Db 457 LYPTIIR 463
:|||||

RESULT 9

F44466
DNA-directed RNA polymerase (EC 2.7.7.6) beta chain - Thermotoga maritima (strain MSB8)
C:Species: Thermotoga maritima
C:Date: 10-Jun-1993 #sequence_revision 26-Apr-1996 #text_change 21-Jul-2000
C:Accession: S41466; S41462; F44466; B72373; S19903
R:Palm, P.; Schleper, C.; Arnold-Ammer, I.; Holz, I.; Meier, T.; Lottspeich, F.; Zillig, W.
Nucleic Acids Res. 21, 4904-4908, 1993
A:Title: The DNA-dependent RNA-polymerase of Thermotoga maritima; characterisation of the

A:Reference number: S41462; MUID:94232816
A:Accession: S41466

A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA

A:Residues: 1-1263 <PAL>

A:Cross-references: EMBL:X72695; NID:g425255; PIDN:CAA51246.1; PID:g425258

A:Experimental source: DSM 3109
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1993

A:Accession: S41462

A:Molecule type: protein

A:Residues: 1-5',X',7-23 <PAW>

A:Experimental source: DSM 3109

R:Liao, D.; Dennis, P.P.
J. Biol. Chem. 267, 22787-22797, 1992

A:Title: The organization and expression of essential transcription translation components of the bacterium Escherichia coli
A:Reference number: A44466; MUID:93054590

A:Accession: F44466

A:Molecule type: DNA

A:Residues: 1-404 <LIA>

A:Cross-references: EMBL:Z11839; NID:g407020; PIDN:CAA77863.1; PID:g48188

A>Note: sequence extracted from NCBI backbone (NCBIP:118059)

R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey, J.R.; Venter, A.; White, O.; Adams, M.D.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, C.M.

Nature 399, 323-329, 1999

A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome

A:Reference number: A72200; MUID:99287316

A:Accession: B72373

A:Molecule type: DNA

A:Residues: 1-1263 <ARN>

A:Cross-references: GB:AE001724; GB:AE000512; NID:g4980966; PIDN:AAD35543.1; PID:g4980966

A:Experimental source: strain MSB8

C:Genetics:

A:Gene: rpoB

C:Superfamily: DNA-directed RNA polymerase beta chain

C:Keywords: nucleotidyltransferase; transcription

F:1-1263/Product: DNA-directed RNA polymerase beta chain #status predicted <MAT>

Query Match 78.0%; Score 32; DB 2; Length 1263;

Best Local Similarity 83.3%; Pred. No. 1.8e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYL 7

Db 214 LYPTYL 219
:|||||

RESULT 10

B85515

hypothetical protein 20321 [imported] - Escherichia coli (strain O157:H7)

C:Species: Escherichia coli

C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 31-Mar-2001

C:Accession: B85515

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May

Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apoda

Nature 408, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: B85515

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-247 <STO>

A:Cross-references: GB:AE005174; NID:gl2513034; PIDN:AAG54582.1; GSPDB:GN00145; UWGP:

A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:

A:Gene: 20321

C:Superfamily: hypothetical protein f242a

Query Match 75.6%; Score 31; DB 2; Length 247;

Best Local Similarity 83.3%; Pred. No. 54;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 YPTYLK 8
 Db 76 YPSYLK 81

RESULT 11

T99698
 Hypothetical protein F31A3.2 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
 C:Accession: T99698
 R:Muray, J.; Le-T.
 Submitted to the EMBL Data Library, May 1996
 A:Description: The sequence of C. elegans cosmid F31A3.
 A:Reference number: Z20667
 A:Accession: T99698
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-262 <MUR>
 A:Cross-references: EMBL:U58742; PIDN:AAB36855.1; GSPDB:GN00028; CESP:F31A3.2
 A:Experimental source: strain Bristol N2; clone F31A3
 C:Genetics:
 A:Gene: CESP:F31A3.2
 A:Map position: X
 A:Introns: 23/3; 80/2; 155/2

Query Match 75.6%; Score 31; DB 2; Length 362;
 Best Local Similarity 71.4%; Pred. No. 58;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 MYPTLK 8
 Db 188 MYPVLE 194

RESULT 12

KCHUM
 Matrilysin (EC 3.4.24.23) precursor - human
 N:Alternate names: matrin; matrix metalloproteinase 7 (MMP7); probable metalloproteinase
 N:Contains: promatrilysin
 C:Species: Homo sapiens (man)
 C>Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 18-Jun-1999
 C:Accession: B28816; A60539; S24324
 R:Muller, D.; Quantin, B.; Gesnel, M.C.; Millon-Collard, R.; Abecassis, J.; Breathnach, Biochem. J. 253, 187-192, 1988
 A:Title: The collagenase gene family in humans consists of at least four members.
 A:Reference number: A90339; MUID:86339885
 A:Accession: B28816
 A:Molecule type: mRNA
 A:Residues: 1-267 <MUL>
 A:Cross-references: EMBL:X07819; NID:g35798; PIDN:CAA30678.1; PID:g35799
 R:Miyazaki, K.; Hattori, Y.; Umenishi, F.; Yasumitsu, H.; Umeda, M.
 Cancer Res. 50, 7758-7764, 1990
 A:Title: Purification and characterization of extracellular matrix-degrading metalloproteinase
 A:Reference number: A60539; MUID:91070531
 A:Accession: A60539
 A:Molecule type: protein
 A:Residues: 18-35, 'x', 37-42 <MIY>
 R:Marti, H.P.; McNeil, L.; Thomas, G.; Davies, M.; Lovett, D.H.
 Biochem. J. 285, 899-905, 1992
 A:Title: Molecular characterization of a low-molecular-mass matrix metalloproteinase secreted by human melanocytes
 A:Reference number: S24324; MUID:92359961
 A:Accession: S24324
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-267 <MAR>
 A:Cross-references: EMBL:Z11887; NID:g35802; PIDN:CAA77942.1; PID:g35803
 C:Comment: This enzyme is similar in its activity to stromelysin and degrades various extracellular matrix components of types II, IV, IX, X, and XI.
 C:Comment: Matrilysin hydrolyzes peptide bonds in plasminogen to yield a fragment with a C-terminal lysine.
 C:Genetics:

A:Gene: GDB:MMP7; MPSL1
 A:Cross-references: GDB:125751; OMIM:178990
 A:Map position: 11q21-11q22
 C:Superfamily: matrilysin; matrix metalloproteinase homology
 C:Keywords: Calcium; extracellular matrix; fibroblast; hydrolase; metalloproteinase;
 F:1-17/Domain: signal sequence #status predicted <SIG>
 F:18-267/Product: promatrilysin #status predicted <PRO>
 F:18-94/Domain: activation peptide #status predicted <ACT>
 F:83-92/Region: autoinhibitory
 F:95-267/Product: matrilysin #status predicted <MAT>
 F:87,214,218,224/Binding site: zinc, catalytic (Cys, His, His, His) (inhibited) #stat
 F:214,218,224/Binding site: zinc, catalytic (His) (active) #status predicted
 F:215/Active site: Glu #status predicted

Query Match 75.6%; Score 31; DB 1; Length 267;
 Best Local Similarity 100.0%; Pred. No. 59;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
 Db 232 MYPTY 236

RESULT 13

A57490
 Matrilysin (EC 3.4.24.23) precursor - rat
 N:Alternate names: matrix metalloproteinase 7 (MMP7)
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 08-Dec-1995 #sequence_revision 08-Dec-1995 #text_change 22-Jun-1999
 C:Accession: A57490
 R:Abramson, S.R.; Conner, G.E.; Nagase, H.; Neuhaus, I.; Woessner Jr., J.F.
 J. Biol. Chem. 270, 16016-16022, 1995
 A:Title: Characterization of rat uterine matrilysin and its cDNA. Relationship to human
 A:Reference number: A57490; MUID:953332299
 A:Accession: A57490
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-267 <ABR>
 A:Cross-references: GB:L24374; NID:g402492; PIDN:AAA9432.1; PID:g402493
 C:Superfamily: matrilysin; matrix metalloproteinase homology
 C:Keywords: hydrolase; metalloproteinase; zinc; zymogen
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-267/Product: matrilysin #status predicted <MAT>
 F:58-262/Domain: matrix metalloproteinase homology <MMP>
 F:90,217,221,227/Binding site: zinc, catalytic (Cys, His, His, His) (inhibited) #stat
 F:217,221,227/Binding site: zinc, catalytic (His) (active) #status predicted
 F:218/Active site: Glu #status predicted

Query Match 75.6%; Score 31; DB 2; Length 267;
 Best Local Similarity 100.0%; Pred. No. 59;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
 Db 235 MYPTY 239

RESULT 14

SEC14 protein - yeast (Candida albicans)
 C:Species: Candida albicans
 C>Date: 13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 26-Feb-1998
 C:Accession: S57923; S72193
 R:Monteoliva, L.; Sanchez, M.; Pla, J.; Gil, C.; Nombela, C.
 Submitted to the EMBL Data Library, September 1994
 A:Description: Characterisation of the Candida albicans SEC14 homolog gene.
 A:Reference number: S57923
 A:Accession: S57923
 A:Status: preliminary
 A:Molecule type: DNA

A:Residues: 1-301 <MON>
 A:Cross-references: EMBL:X81937
 R:Montecolliva, L.; Sanchez, M.; Pla, J.; Gil, C.; Nombela, C.
 Yeast 12, 1097-1105, 1996
 A:Title: Cloning of Candida albicans SEC14 gene homologue coding for a putative essential protein
 A:Reference number: S72193; MUID:97051600
 A:Accession: S72193
 A:Molecule type: DNA
 A:Residues: 1-301 <MON>
 A:Cross-references: EMBL:X81937
 A:Note: the authors translated the codon CTG for residue 180 as Ser
 C:Genetics:
 A:Gene: SEC14
 C:Superfamily: cellular retinaldehyde-binding protein; cellular retinaldehyde-binding protein
 F:59-266/Domain: cellular retinaldehyde-binding protein homology <CRB>

Query Match 75.6%; Score 31; DB 2; Length 301;
 Best Local Similarity 100.0%; Pred. No. 66;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2 MYPTY 6
 |||||
 Db 106 MYPTY 110

RESULT 15
 F86307
 hypothetical protein AAD50017.1 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence-revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: F86307
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719
 A:Accession: F86307
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-378 <STO>
 A:Cross-references: GB:AE005172; NID:g5734752; PIDN:AAD50017.1; GSPDB:GN00141
 C:Genetics:
 Map position: 1

Query Match 75.6%; Score 31; DB 2; Length 378;
 Best Local Similarity 100.0%; Pred. No. 84;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
 |||||
 Db 118 MYPTY 122

Search completed: September 4, 2001, 15:52:09
 Job time: 232 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2001, 15:54:10 ; Search time 27.19 Seconds
(without alignments)
10.079 Million cell updates/sec

Title: US-09-630-345-6
Perfect score: 41
Sequence: 1 XMPYLYK 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	37	90.2	384	MM01_RANCA	Q11133 rana catesb
2	34	82.9	436	GABT_BACSU	P94427 bacillus su
3	32	78.0	340	CELF_VZVD	P09261 varicella-2
4	32	78.0	349	PERE_ARATH	P24102 arabidopsis
5	32	78.0	594	NUSM_HIPAM	Q9ZZY1 hippopotamu
6	32	78.0	1263	RPOB_THEMA	P29398 thermotoga
7	31	75.6	262	MM07_FELCA	P55032 felis silve
8	31	75.6	264	MM07_MOUSE	Q10738 mus musculu
9	31	75.6	267	MM07_HUMAN	P09237 homo sapien
10	31	75.6	267	MM07_RAT	P50280 rattus norv
11	31	75.6	301	SC14_CANAL	P46250 candida alb
12	31	75.6	462	MM12_MOUSE	P34960 mus musculu
13	31	75.6	465	MM12_RAT	Q63341 rattus norv
14	31	75.6	467	MM18_XENLA	Q13065 xenopus lae
15	31	75.6	482	MM20_MOUSE	P57748 mus musculu
16	31	75.6	483	MM20_HUMAN	Q60882 homo sapien
17	31	75.6	483	MM20_PIG	P79287 sus scrofa
18	31	75.6	495	VLI1_BPVI	P03103 bovine papi
19	31	75.6	497	VLI1_BPVI	P06458 bovine papi
20	31	75.6	501	VLI1_PAPVE	P11326 european el
21	31	75.6	503	YPTJ_CAEEL	Q23459 caenorhabdi
22	31	75.6	503	TUB_MOUSE	P50586 mus musculu
23	31	75.6	505	TUB_RAT	Q88808 rattus norv
24	31	75.6	506	TUB_HUMAN	P50607 homo sapien
25	31	75.6	511	YELJ_CAEEL	P90756 caenorhabdi
26	31	75.6	513	VLI1_PAPVD	P03104 deer papill
27	31	75.6	662	YMEI_SCHMA	P46508 schistosoma
28	31	75.6	2185	POLG_SVDVH	P16604 s genome po
29	31	75.6	2185	POLG_SVDVU	P13900 s genome po
30	30	73.2	91	Y13F_BPT4	P39499 bacterioph
31	30	73.2	192	NUGM_PROWI	Q37622 prototheca
32	30	73.2	239	YKGE_ECOLI	P77252 escherichia
33	30	73.2	308	Y040_BPT4	P39254 bacterioph

34 30 73.2 343 1 CELF_HSVB P28933 equine herp
35 30 73.2 368 1 VGH3_HCMVA P08560 human cytom
36 30 73.2 425 1 BIOA_SERMA P36568 serratia ma
37 30 73.2 580 1 P69_MYCHR P15362 mycoplasma
38 30 73.2 634 1 VC09_VACCC P21042 vaccinia v1
39 30 73.2 634 1 VC09_VACCV P17372 vaccinia v1
40 30 73.2 1056 1 DPOL_ADE02 P03261 human adeno
41 30 73.2 1056 1 DPOL_ADE05 P0495 human adeno
42 30 73.2 1138 1 C7AA_BACTU Q03749 bacillus th
43 30 73.2 1138 1 C7AB_BACUA Q45707 bacillus th
44 30 73.2 1138 1 C7AB_BACUK Q45708 bacillus th
45 30 73.2 1155 1 C1AB_BACTK P06578 bacillus th

ALIGNMENTS

RESULT 1
MM01_RANCA STANDARD; PRT; 384 AA.
ID Q11133;
AC 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE INTERSTITIAL COLLAGENASE PRECURSOR (EC 3.4.24.7) (MATRIX
DE METALLOPROTEINASE-1) (MMP-1) (TCL).
OS Rana catesbeiana (Bull frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Rana.
OX NCBI_TaxID=8400;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RX MEDLINE=95071832; PubMed=7981043;
RA Oofusa K., Yomori S., Yoshizato K.;
RT Regionally and hormonally regulated expression of genes of collagen
RT and collagenase in the anuran larval skin.";
RL Int. J. Dev. Biol. 38:345-350(1994).
CC -!- FUNCTION: CLEAVES COLLAGENS OF TYPES I, II, AND III AT ONE SITE IN
CC THE HELICAL DOMAIN. ALSO CLEAVES COLLAGENS OF TYPES VII AND X.
CC -!- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY.
CC -!- ENZYME REGULATION: CAN BE ACTIVATED WITHOUT REMOVAL OF THE
CC ACTIVATION PEPTIDE (BY SIMILARITY).
CC -!- SIMILARITY: CONTAINS 1 HEMOPEXIN-LIKE DOMAIN.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC
CC METALLOPROTEASE) ALSO KNOWN AS MATRIXIN SUBFAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; S75623; AAB32661.1; -;
CC HSSP; P21692; IFBL.
CC MEROPS; M10.001; -;
CC InterPro; IPR000130; -;
CC InterPro; IPR000585; -;
CC InterPro; IPR001818; -;
CC Pfam; PF00413; Peptidase_M10; 1.
CC Pfam; PF00045; hemopexin; 2.
CC PROSITE; PS00024; HEMOPEXIN; FALSE_NEG.
CC PROSITE; PS00142; ZINC_PROTEASE; 1.
CC PROSITE; PS00546; CYSTEINE_SWITCH; FALSE_NEG.
CC Hydrolase; Metalloprotease; Zinc; Zymogen; Calcium;
CC Collagen degradation; Extracellular matrix; Signal.
CC SIGNAL 1 25 POTENTIAL.
CC PROPEP 26 88 ACTIVATION PEPTIDE (POTENTIAL).
CC FT CHAIN 29 384 INTERSTITIAL COLLAGENASE.
CC FT DOMAIN 83 384 HEMOPEXIN-LIKE.
CC SITE 81 81 CYSTEINE SWITCH (POTENTIAL).

FT METAL 189 189 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACT_SITE 190 190 BY SIMILARITY.
 FT METAL 193 193 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 199 199 ZINC (CATALYTIC) (BY SIMILARITY).
 FT DISULFID 242 381 PROBABLE.
 SQ SEQUENCE 384 AA; 43582 MW; A5B5E2FB32239DF CRC64;
 Query Match 90.2%; Score 37; DB 1; Length 384;
 Best Local Similarity 85.7%; Pred. No. 2.9;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 2 MYPTYLK 8
 DB 207 MYPTYLK 213
 RESULT 2
 GABT_BACSU STANDARD; PRT; 436 AA.
 P94427;
 15-DEC-1998 (Rel. 37, Created)
 30-MAY-2000 (Rel. 39, Last annotation update)
 DE PROBABLE 4-AMINO BUTYRATE AMINOTRANSFERASE (EC 2.6.1.19) (GAMMA-AMINO-
 DE N-BUTYRATE TRANSAMINASE) (GABA TRANSAMINASE) (GLUTAMATE:SUCCINIC
 DE SEMIALDEHYDE TRANSAMINASE) (GABA AMINOTRANSFERASE) (GABA-AT).
 GN GABT.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=97124189; PubMed=8969502;
 RA Yamane K., Kumano M., Kurita K.;
 RT "The 25 degrees-36 degrees region of the Bacillus subtilis
 RT chromosome: determination of the sequence of a 146 kb segment and
 RT identification of 113 genes."
 RL Microbiology 142:3047-3056(1996).
 CC -|- CATALYTIC ACTIVITY: 4-AMINOBUTANOATE + 2-OXOGLUTARATE -> SUCCINATE
 CC -|- SEMIALDEHYDE + L-GLUTAMATE.
 CC -|- COFACTOR: PYRIDOXAL PHOSPHATE.
 CC -|- PATHWAY: 4-AMINO BUTYRATE (GABA) DEGRADATION PATHWAY.
 CC -|- SIMILARITY: BELONGS TO CLASS-III OF PYRIDOXAL-PHOSPHATE-DEPENDENT
 CC AMINOTRANSFERASES.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).

DR EMBL; D50453; BAA03021.1;
 DR EMBL; 299106; BAB12198.1;
 DR Subtilist; BG12043; gabt.
 DR InterPro; IPR000954;
 DR Pfam; PF00202; aminotran_3.1.
 DR PROSITE; PS00600; AA_TRANSF_0600.
 KW Transferrase; Aminotransferase; Pyridoxal phosphate.
 FT BINDING 281 281 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 SQ SEQUENCE 436 AA; 47249 MW; D0961F6D4189A8F3 CRC64;

Query Match 82.9%; Score 34; DB 1; Length 436;
 Best Local Similarity 71.4%; Pred. No. 13;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
 DB 207 MYPTYLK 213

DB 88 MYPTYLK 94
 RESULT 3
 CELF_VZVD STANDARD; PRT; 340 AA.
 ID CELF_VZVD
 AC P09261;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE CELL FUSION PROTEIN PRECURSOR.
 GN 5
 OS Varicella-zoster virus (strain Dumas) (VZV).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirus.
 OX NCBI_TaxID=10338;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86306657; PubMed=3018124;
 RA Davison A.J., Scott J.E.;
 RT "The complete DNA sequence of varicella-zoster virus."
 RL J. Gen. Virol. 67:1759-1816(1986).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X04370; CAA27888.1;
 DR PIR; E27212; MMBE5
 DR InterPro; IPR002567;
 DR Pfam; PF01621; Fusion_gly_K; 1.
 KW Fusion protein; Transmembrane; Signal.
 FT SIGNAL 1 7
 FT CHAIN 1 340 CELL FUSION PROTEIN.
 FT CARBOHYD 55 55 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 65 65 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 340 AA; 38576 MW; 0387FE0EC39C946 CRC64;
 Query Match 78.0%; Score 32; DB 1; Length 340;
 Best Local Similarity 57.1%; Pred. No. 25;
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 2 MYPTYLK 8
 DB 244 LYPTVIR 250
 RESULT 4
 PERE_ARATH STANDARD; PRT; 349 AA.
 ID PERE_ARATH
 AC P24102;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE BASIC PEROXIDASE E PRECURSOR (EC 1.11.1.7).
 GN PRXA.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91200671; PubMed=2016063;
 RA Intaprak C., Higashimura N., Yamamoto K., Okada N., Shimmyo A.,
 RA Takano M.;
 RT "Nucleotide sequences of two genomic DNAs encoding peroxidase of
 Arabidopsis thaliana."

```

RL Gene 98:237-241(1991).
CC -1- FUNCTION: REMOVAL OF H(2)O(2), OXIDATION OF TOXIC REDUCTANTS,
CC BIOSYNTHESIS AND DEGRADATION OF LIGNIN, DEFENSE RESPONSE TOWARD
CC WOUNDING AND METABOLISM OF AUXIN. THESE FUNCTIONS MIGHT BE
CC DEPENDENT ON EACH ISOZYME/ISOFORM IN EACH PLANT TISSUE.
CC -1- CATALYTIC ACTIVITY: DONOR + H(2)O(2) = OXIDIZED DONOR + 2 H(2)O.
CC -1- COFACTOR: HEME.
CC -1- TISSUE SPECIFICITY: ROOTS.
CC -1- FAMILY: BELONGS TO THE PEROXIDASE FAMILY. PLANT PEROXIDASE
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL: M58381; AAA32842.1; -.
PIR: JU0458; JU0458.
HSSP: P00433; IATJ.
DR InterPro: IPR000823; -.
DR InterPro: IPR002016; -.
DR Pfam: PF00141; peroxidase.1.
DR PRINTS: PR00458; PEROXIDASE.
DR PRINTS: PR00461; PLPEROXIDASE.
DR PROSITE: PS00435; PEROXIDASE.1; 1.
DR PROSITE: PS00436; PEROXIDASE.2; 1.
KW Oxidoreductase; Glycoprotein; Peroxidase; Heme; Multigene family;
KW Signal.
FT SIGNAL. 1 19
FT CHAIN 20 349 BASIC PEROXIDASE E.
FT ACT_SITE 67 67 BY SIMILARITY.
FT ACT_SITE 71 71 DISTAL HISTIDINE (BY SIMILARITY).
FT ACT_SITE 199 199 PROXIMAL HISTIDINE (HEME AXIAL LIGAND).
FT DISULFID 40 120 BY SIMILARITY.
FT DISULFID 73 78 BY SIMILARITY.
FT DISULFID 126 329 BY SIMILARITY.
FT DISULFID 206 238 BY SIMILARITY.
FT CARBOHYD 86 86 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 173 173 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 187 187 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 217 217 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 243 243 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 349 AA; 38172 MW; 6E02C5758C36AB8F CRC64;

Query Match 78.0%; Score 32; DB 1; Length 349;
Best Local Similarity 83.3%; Pred. No. 26;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYL 7
Db 226 LYPTYL 231
:|||||

RESULT 5
NU5M_HIPAM STANDARD; PRT; 594 AA.
AC Q92ZV1.
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 5 (EC 1.6.5.3).
GN MTND5 OR ND5 OR NADH5.
OS Hippopotamus amphibius (Hippopotamus).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Hippopotamidae; Hippopotamus.
OX NCBI_TaxID=9833;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RA Ursing B.M., Arnason U.;
RT "Analyses of mitochondrial genomes strongly support a hippopotamus-
RT whale clade."
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL: AJ010957; CAA09438.1; -.
DR InterPro: IPR001516; -.
DR InterPro: IPR001750; -.
DR Pfam: PF00361; oxidored_g1.1.
DR Pfam: PF00662; oxidored_g1_N; 1.
KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
SQ SEQUENCE 594 AA; 66599 MW; 4130499096B5A5CE CRC64;

Query Match 78.0%; Score 32; DB 1; Length 594;
Best Local Similarity 83.3%; Pred. No. 44;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 YPTYLK 8
Db 32 YPTYVK 37
|||||

RESULT 6
RPOB_THEMA STANDARD; PRT; 1263 AA.
ID RPOB_THEMA
AC P29398.
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE DNA-DIRECTED RNA POLYMERASE BETA CHAIN (EC 2.7.7.6) (TRANSCRIPTASE
DE BETA CHAIN) (RNA POLYMERASE BETA SUBUNIT).
GN RPOB OR TM0458.
OS Thermotoga maritima.
OC Bacteria; Thermotogales; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109;
RX MEDLINE=94232816; PubMed=8177738;
RA Palm P., Schleper C., Arnold-Ammer I., Holz I., Meier T.,
RA Lottspeich F., Zillig W.;
RT "The DNA-dependent RNA-polymerase of Thermotoga maritima;
RT characterisation of the enzyme and the DNA-sequence of the genes for
RT the large subunits."
RL Nucleic Acids Res. 21:4904-4908(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
RT genome sequence of Thermotoga maritima."
RL Nature 399:323-329(1999).
RN [3]
RP SEQUENCE OF 1-404 FROM N.A.
RC STRAIN=MSB8 / DSM 3109;
RX MEDLINE=93054590; PubMed=1429627;
RA Liao D., Dennis P.P.;

```

RT The organization and expression of essential transcription
 RT translation component genes in the extremely thermophilic eubacterium
 RL Thermotoga maritima.";
 RL J. Biol. Chem. 267:22787-22797(1992).
 CC -!- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
 CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
 CC SUBSTRATES.
 CC -!- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE + N PYROPHOSPHATE +
 CC RNA(N).
 CC -!- SUBUNIT: THE ENZYME CONSISTS OF THE SIGMA CHAIN AND THE CORE
 CC ENZYME WHICH IS COMPOSED OF 2 ALPHA CHAINS, 1 BETA CHAIN, AND 1
 CC BETA' CHAIN.
 CC -!- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA CHAIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X72695; CAA51246.1; -
 CC EMBL; AE001724; AAD35543.1; -
 CC EMBL; Z11839; CAA77863.1; -
 CC PIR; S19903; S19903.
 CC PIR; S41466; S41466.
 CC TIGR; TM0458; -
 CC InterPro; IPR001572; -
 CC Pfam; PF00562; RNA_pol_B; 1.
 CC PROSITE; PS01166; RNA_POL.BETA.1.
 KW Transferrase; Transcription; DNA-directed RNA polymerase.
 SQ SEQUENCE 1263 AA; 143137 MW; 04B79368567C8237 CRC64;

Query Match 78.0%; Score 32; DB 1; Length 1263;
 Best Local Similarity 83.3%; Pred. No. 94;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPYL 7
 DB 214 LYPTIL 219

RESULT 7
 MM07_FELCA STANDARD; PRT; 262 AA.
 ID MM07_FELCA
 AC P55032;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE MATRILYSIN PRECURSOR (EC 3.4.24.23) (PUMP-1 PROTEASE) (UTERINE
 DE METALLOPROTEINASE) (MATRIX METALLOPROTEINASE-7) (MMP-7) (MATRIN)
 DE (FRAGMENT).
 DE MMP7.
 GN Felis silvestris catus (Cat).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
 OC NCBI_TaxID=9685;
 OX [1]
 RN TISSUE=Endometrium;
 RP SEQUENCE FROM N.A.
 RC Scelzo C.M., Verhage H.G., Jaffe R.C.;
 RA "Expression and estrogen control of PUMP-1 mRNA in the cat uterus.";
 RL Endocrinol. Jpn. 2:229-235(1994).
 CC -!- FUNCTION: DEGRADES CASEIN, GELATINS OF TYPES I, III, IV, AND V,
 CC AND FIBRONECTIN. ACTIVATES PROCOLLAGENASE.
 CC -!- CATALYTIC ACTIVITY: CLEAVAGE OF 14-ALA-|-LEU-15 AND 16-TYR-|-
 CC LEU-17 IN B CHAIN OF INSULIN. NO ACTION ON COLLAGEN TYPES I, II,
 CC IV. V. CLEAVES GELATIN CHAIN ALPHA-2(I) > ALPHA-1(I).
 CC -!- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC
 CC METALLOPROTEASE) ALSO KNOWN AS MATRININ SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U04444; AAA18222.1; -
 CC HSSP; P09237; LMMR.
 CC MEROPS; M10.008; -
 CC InterPro; IPR000130; -
 CC InterPro; IPR0001818; -
 CC Pfam; PF04113; Peptidase_M10; 1.
 CC PROSITE; PS00142; CYSTEINE_SWITCH; 1.
 CC PROSITE; PS00546; CYSTEINE_SWITCH; 1.
 KW Hydrolase; Metalloprotease; Zinc; Zymogen; Calcium;
 KW Collagen degradation; Extracellular matrix; Signal.
 FT NON_TER 1
 FT SIGNAL <1 12
 FT PROPEP 13 89
 FT CHAIN 90 262
 FT SITE 82 82
 FT METAL 209 209
 FT ACT_SITE 210 210
 FT METAL 213 213
 FT METAL 219 219
 FT SIGNAL 262 262
 SQ SEQUENCE 262 AA; 29263 MW; B4A1FA23320DC732 CRC64;
 BY SIMILARITY.
 ACTIVATION PEPTIDE (BY SIMILARITY).
 MATRILYSIN.
 CYSTEINE SWITCH (POTENTIAL).
 ZINC (CATALYTIC) (BY SIMILARITY).
 BY SIMILARITY.
 ZINC (CATALYTIC) (BY SIMILARITY).
 ZINC (CATALYTIC) (BY SIMILARITY).
 ZINC (CATALYTIC) (BY SIMILARITY).

Query Match 75.6%; Score 31; DB 1; Length 262;
 Best Local Similarity 100.0%; Pred. No. 31;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPYT 6
 DB 227 MYPYT 231

RESULT 8
 MM07_MOUSE STANDARD; PRT; 264 AA.
 ID MM07_MOUSE
 AC Q10738;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE MATRILYSIN PRECURSOR (EC 3.4.24.23) (PUMP-1 PROTEASE) (UTERINE
 DE METALLOPROTEINASE) (MATRIX METALLOPROTEINASE-7) (MMP-7) (MATRIN).
 DE MMP7.
 GN Mus musculus (Mouse).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 OX [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=ICE; TISSUE=Uterus;
 RA MEDLINE=96086498; PubMed=7579699;
 RA Wilson C.L., Heppner K.J., Rudolph L.A., Matrisian L.M.;
 RT "The metalloproteinase matrilysin is preferentially expressed by
 RT epithelial cells in a tissue-restricted pattern in the mouse.";
 RL Mol. Biol. Cell 6:851-869(1995).
 CC -!- FUNCTION: DEGRADES CASEIN, GELATINS OF TYPES I, III, IV, AND V,
 CC AND FIBRONECTIN. ACTIVATES PROCOLLAGENASE (BY SIMILARITY).
 CC -!- FUNCTION: MAY PLAY A ROLE IN TISSUE REORGANIZATION.
 CC -!- CATALYTIC ACTIVITY: CLEAVAGE OF 14-ALA-|-LEU-15 AND 16-TYR-|-
 CC LEU-17 IN B CHAIN OF INSULIN. NO ACTION ON COLLAGEN TYPES I, II,
 CC IV. V. CLEAVES GELATIN CHAIN ALPHA-2(I) > ALPHA-1(I).
 CC -!- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC
 CC METALLOPROTEASE) ALSO KNOWN AS MATRININ SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch)
CC -----
DR EMBL; L36238; AAA99984.1; JOINED.
DR EMBL; L36243; AAA99984.1; JOINED.
DR EMBL; L36242; AAA99984.1; JOINED.
DR EMBL; L36241; AAA99984.1; JOINED.
DR EMBL; L36240; AAA99984.1; JOINED.
DR EMBL; L36239; AAA99984.1; JOINED.
DR EMBL; L36244; AAA99983.1; JOINED.
DR HSP; P09237; IMMR.
DR MEROPS; M10.008; -.
DR MGD; MGI:103189; Mmp7.
DR InterPro; IPR000130; -.
DR InterPro; IPR001818; -.
DR Pfam; PF00413; Peptidase_M10; 1.
DR PRINTS; PR00138; MATRIXIN.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
DR PROSITE; PS00546; CYSTEINE_SWITCH; 1.
DR Hydrolase; Metalloprotease; Zinc; Zymogen; Calcium;
KW Collagen degradation; Extracellular matrix; Signal.
FT SIGNAL 1 17 POTENTIAL.
FT PROPEP 18 94 ACTIVATION PEPTIDE (BY SIMILARITY).
FT CHAIN 95 264 MATRILYSIN.
FT SITE 87 87 CYSTEINE SWITCH (POTENTIAL).
FT METAL 214 214 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 215 215 BY SIMILARITY.
FT METAL 218 215 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 224 224 ZINC (CATALYTIC) (BY SIMILARITY).
FT CONFLICT 201 201 G -> D (IN AAA99983).
SQ SEQUENCE 264 AA; 29755 MW; EDA31A5EBAC63342 CRC64;

Query Match 75.6%; Score 31; DB 1; Length 264;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPT 6
IIIIII
Db 232 MYPT 236

RESULT 9
MM07_HUMAN STANDARD; PRT; 267 AA.
P09237;
01-MAR-1989 (Rel. 10, Created)
01-MAR-1989 (Rel. 10, Last sequence update)
01-OCT-2000 (Rel. 40, Last annotation update)
DE MATRILYSIN PRECURSOR (EC 3.4.24.23) (PUMP-1 PROTEASE) (UTERINE
DE METALLOPROTEINASE) (MATRIX METALLOPROTEINASE-7) (MMP-7) (MATRIN).
GN MMP7 OR MPSL1 OR PUMPL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88339885; PubMed=2844164;
RA Muller D., Quantin B., Gesnel M.-C., Millon-Collard R., Abecassis J.,
RA Breathnach R.;
RT "The collagenase gene family in humans consists of at least four
RT members.";
RL Biochem. J. 253:187-192(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=9235961; PubMed=1497627;
RA Marti H.P., McNeill L., Thomas G., Davies M., Lovett D.H.;

RT "Molecular characterization of a low-molecular-mass matrix
RT metalloproteinase secreted by glomerular mesangial cells as PUMP-1.";
RL Biochem. J. 285:899-905(1992).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=94124555; PubMed=8294454;
RA Gaire M., Magbanua Z., McDonnell S., McNeill L.B., Lovett D.H.,
RA Matristian L.M.;
RT "Structure and expression of the human gene for the matrix
RT metalloproteinase matrilysin.";
RL J. Biol. Chem. 269:2032-2040(1994).
RN [4]
RP SEQUENCE OF 18-42.
RX MEDLINE=91070531; PubMed=2253219;
RA Miyazaki K., Hattori Y., Umenishi F., Yasumitsu H., Umeda M.;
RT "Purification and characterization of extracellular matrix-degrading
RT metalloproteinase, matrin (pump-1), secreted from human rectal
RT carcinoma cell line.";
RL Cancer Res. 50:7758-7764(1990).
RN [5]
RP FUNCTION.
RX MEDLINE=89375247; PubMed=2550050;
RA Quantin B., Murphy G., Breathnach R.;
RT "Pump-1 cDNA codes for a protein with characteristics similar to
RT those of classical collagenase family members.";
RL Biochemistry 28:5327-5334(1989).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RX MEDLINE=9575856; PubMed=7756291;
RA Browner M.F., Smith W.W., Castelhan A.L.;
RT "Matrilysin-inhibitor complexes: common themes among
RT metalloproteases.";
RL Biochemistry 34:6602-6610(1995).
CC -!- FUNCTION: DEGRADATES CASEIN, GELATINS OF TYPES I, III, IV, AND V,
CC AND FIBRONECTIN. ACTIVATES PROCOLLAGENASE.
CC -!- CATALYTIC ACTIVITY: CLEAVAGE OF 14-ALA-1-LEU-15 AND 16-TYR-1-
CC LEU-17 IN B CHAIN OF INSULIN. NO ACTION ON COLLAGEN TYPES I, II,
CC IV, V. CLEAVES GELATIN CHAIN ALPHA-2(1) > ALPHA-1(1).
CC -!- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC
CC METALLOPROTEASE) ALSO KNOWN AS MATRININ SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch)
CC -----
DR EMBL; X07819; CAA30678.1; -.
DR EMBL; Z11887; CAA77942.1; -.
DR EMBL; L22524; AAC37543.1; -.
DR EMBL; L22519; AAC37543.1; JOINED.
DR EMBL; L22520; AAC37543.1; JOINED.
DR EMBL; L22521; AAC37543.1; JOINED.
DR EMBL; L22522; AAC37543.1; JOINED.
DR EMBL; L22523; AAC37543.1; JOINED.
DR PIR; B28816; KCHUM.
DR PIR; S24324; S24324.
DR PDB; 1MMQ; 03-APR-96.
DR PDB; 1MMR; 03-APR-96.
DR MEROPS; M10.008; -.
DR MIM; I78990; -.
DR InterPro; IPR000130; -.
DR InterPro; IPR001818; -.
DR Pfam; PF00413; Peptidase_M10; 1.
DR PRINTS; PR00138; MATRIXIN.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
DR PROSITE; PS00546; CYSTEINE_SWITCH; 1.
KW Hydrolase; Metalloprotease; Zinc; Zymogen; Calcium; 3D-structure;
KW

Query Match 75.6%; Score 31; DB 1; Length 301;
 Best Local Similarity 100.0%; Pred. No. 36;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
 |||||
 DB 106 MYPTY 110

RESULT 12
 MM12_MOUSE STANDARD; PRT; 462 AA.
 AC P34960;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE MACROPHAGE METALLOELASTASE PRECURSOR (EC 3.4.24.65) (MME) (MATRIX
 METALLOPROTEINASE-12) (MMP-12).
 MM12 OR MMEL OR MME.
 Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 99-125.
 RC TISSUE=Macrophage;
 RX MEDLINE=92165826; PubMed=1537850;
 RA Shapiro S.D., Griffin G.L., Gilbert D.J., Jenkins N.A.,
 RA Copeland N.G., Welgus H.G., Senior R.M., Ley T.J.;
 RT "Molecular cloning, chromosomal localization, and bacterial
 expression of a murine macrophage metalloelastase.";
 RL J. Biol. Chem. 267:4664-4671(1992).
 CC -1- FUNCTION: MAY BE INVOLVED IN TISSUE INJURY AND REMODELING. HAS
 CC -1- SIGNIFICANT ELASTOLYTIC ACTIVITY.
 CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF SOLUBLE AND INSOLUBLE ELASTIN.
 CC SPECIFIC CLEAVAGES ARE ALSO PRODUCED AT 14-ALA-|-LEU-15 AND 16-
 CC TYR-|-LEU-17 IN THE B CHAIN OF INSULIN.
 CC -1- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY.
 CC -1- SIMILARITY: CONTAINS 1 HEMOPEXIN-LIKE DOMAIN.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC
 CC METALLOPROTEASE) ALSO KNOWN AS MATRXIN SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M82831; AAA39526.1; --
 CC FIR; A42401; A42401.
 CC HSP; P03956; ICGL.
 CC MEROPS; M10.009; --
 CC MGD; MGI:97005; Mmp12.
 CC InterPro; IPR000130; --
 CC InterPro; IPR000585; --
 CC InterPro; IPR001818; --
 CC Pfam; PF00413; Peptidase_M10; 1.
 CC Pfam; PF00045; hemopexin; 4.
 CC PRINTS; PR00138; MATRXIN.
 CC PROSITE; PS00024; HEMOPEXIN; 1.
 CC PROSITE; PS00142; ZINC_PROTEASE; 1.
 CC PROSITE; PS00546; CYSTEINE_SWITCH; 1.
 CC Hydrolase; Metalloprotease; Glycoprotein; Zinc; Zymogen; Calcium;
 CC Extracellular matrix; Signal.
 CC SIGNAL 1 17 PROBABLE.
 FT PROPEP 18 98 ACTIVATION PEPTIDE.
 FT CHAIN 99 462 MACROPHAGE METALLOELASTASE.
 FT DOMAIN 272 462 HEMOPEXIN-LIKE.
 FT SITE 85 85 CYSTEINE SWITCH (BY SIMILARITY).

FT METAL 211 211 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACT_SITE 212 212 BY SIMILARITY.
 FT METAL 215 215 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 221 221 ZINC (CATALYTIC) (BY SIMILARITY).
 FT CARBOHYD 21 21 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 74 74 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 310 310 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT DISULFID 275 462 BY SIMILARITY.
 SQ SEQUENCE 462 AA; 53841 MW; BB9625906FIDBEDF CRC64;
 Query Match 75.6%; Score 31; DB 1; Length 462;
 Best Local Similarity 100.0%; Pred. No. 55;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 MYPTY 6
 |||||
 DB 229 MYPTY 233
 RESULT 13
 MM12_RAT STANDARD; PRT; 465 AA.
 ID MM12_RAT
 AC Q63341;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE MACROPHAGE METALLOELASTASE PRECURSOR (EC 3.4.24.65) (MME) (MATRIX
 METALLOPROTEINASE-12) (MMP-12).
 GN MMP12 OR MMEL.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY;
 RA Cossins J., Clements J., Catlin G.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: MAY BE INVOLVED IN TISSUE INJURY AND REMODELING. HAS
 CC -1- SIGNIFICANT ELASTOLYTIC ACTIVITY (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF SOLUBLE AND INSOLUBLE ELASTIN.
 CC SPECIFIC CLEAVAGES ARE ALSO PRODUCED AT 14-ALA-|-LEU-15 AND 16-
 CC TYR-|-LEU-17 IN THE B CHAIN OF INSULIN.
 CC -1- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY (BY SIMILARITY).
 CC -1- SIMILARITY: CONTAINS 1 HEMOPEXIN-LIKE DOMAIN.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC
 CC METALLOPROTEASE) ALSO KNOWN AS MATRXIN SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X98517; CAA67142.1; --
 CC HSP; P03956; ICGL.
 CC MEROPS; M10.009; --
 CC InterPro; IPR000130; --
 CC InterPro; IPR000585; --
 CC InterPro; IPR001818; --
 CC Pfam; PF00413; Peptidase_M10; 1.
 CC Pfam; PF00045; hemopexin; 4.
 CC PRINTS; PR00138; MATRXIN.
 CC PROSITE; PS00024; HEMOPEXIN; 1.
 CC PROSITE; PS00142; ZINC_PROTEASE; 1.
 CC PROSITE; PS00546; CYSTEINE_SWITCH; 1.
 CC Hydrolase; Metalloprotease; Glycoprotein; Zinc; Zymogen; Calcium;
 CC Extracellular matrix; Signal.
 CC SIGNAL 1 21 PROBABLE.
 FT PROPEP 22 101 ACTIVATION PEPTIDE (BY SIMILARITY).

```

FT CHAIN 102 465 MACROPHAGE METALLOELASTASE.
FT DOMAIN 275 465 HEMOPEXIN-LIKE.
FT SITE 88 88 CYSTEINE SWITCH (BY SIMILARITY).
FT METAL 214 215 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 215 215 BY SIMILARITY.
FT METAL 218 218 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 224 224 ZINC (CATALYTIC) (BY SIMILARITY).
FT CARBOHYD 313 313 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 278 465 BY SIMILARITY.
FT SEQUENCE 465 AA; 53738 MW; E779B6014EC6FF68 CRC64;

Query Match 75.6%; Score 31; DB 1; Length 465;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
DB 232 MYPTY 236

-----
RESULT 14
MM18_XENLA STANDARD; PRT; 467 AA.
AC 013065;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DE 01-OCT-2000 (Rel. 40, Last annotation update)
DE MATRIX METALLOPROTEINASE-18 PRECURSOR (EC 3.4.24.-) (MMP-18)
DE (COLLAGENASE-4) (COLLAGENASE 4) (XCOLA4).
DE MMP18 OR COLA4.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xopodidae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue=Intestine;
RX MEDLINE=97053976; PubMed=8998355;
RA Sang O.A., Shi Y.-B.,
RA Sedgwick D.D., Li J., Sedgwick T., Liang V.C.-T.,
RA "Identification and characterization of a novel collagenase in Xenopus
RA laevis: possible roles during frog development.";
RA Mol. Biol. Cell 7:1471-1483(1996).
CC -1- FUNCTION: CLEAVES COLLAGEN TYPE I. MAY PLAY A ROLE IN LARVAL
CC METAMORPHOSIS. MAY BE INVOLVED IN TAIL RESORPTION.
CC TISSUE DEGENERATION AND ADULT ORGANOGENESIS DURING AMPHIBIAN
CC METAMORPHOSIS.
CC -1- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY (BY SIMILARITY).
CC -1- ENZYME REGULATION: UPREGULATED IN THE TAIL BY THYROID HORMONE.
CC -1- SUBCELLULAR LOCATION: SECRETED. ASSOCIATED WITH THE EXTRACELLULAR
CC MATRIX (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: EXPRESSED ONLY TRANSIENTLY IN WHOLE ANIMAL, AT
CC TIME WHEN TADPOLE FEEDING BEGINS.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED AT HIGH LEVEL AS THE TADPOLE TAIL
CC RESORBS AND DURING Hindlimb morphogenesis and intestinal
CC REMODELING.
CC -1- SIMILARITY: CONTAINS 1 HEMOPEXIN-LIKE DOMAIN.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC
CC METALLOPROTEASE) ALSO KNOWN AS MATRININ SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L76275; AAB53148.1;
CC HSSP; P03956; 1CGL.
CC MEROPS; M10.018;
CC InterPro; IPR000130;

```

```

DR InterPro; IPR000585;
DR InterPro; IPR001818;
DR Pfam; PF00045; Hemoexin; 4.
DR Pfam; PF00413; Peptidase_M10; 1.
DR PRINTS; PR00138; MATRININ
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
DR PROSITE; PS00546; CYSTEINE_SWITCH; FALSE_NEG.
DR PROSITE; PS00024; HEMOPEXIN; FALSE_NEG.
DR Hydrolase; Metalloprotease; zinc; Calcium; Zymogen; Signal;
KW Collagen degradation; Extracellular matrix.
FT SIGNAL 1 17 POTENTIAL.
FT PROPEP 18 99
FT CHAIN 100 467 MATRIX METALLOPROTEINASE-18.
FT DOMAIN 277 467 HEMOPEXIN-LIKE.
FT SITE 92 92 CYSTEINE SWITCH (POTENTIAL).
FT METAL 218 218 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 219 219 BY SIMILARITY.
FT METAL 222 222 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 228 228 ZINC (CATALYTIC) (BY SIMILARITY).
FT DISULFID 280 467 POTENTIAL.
FT SEQUENCE 467 AA; 52812 MW; 4623F6CF3454051 CRC64;

Query Match 75.6%; Score 31; DB 1; Length 467;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTY 6
DB 236 MYPTY 240

-----
RESULT 15
MM20_MOUSE STANDARD; PRT; 482 AA.
AC P57748;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DE 01-OCT-2000 (Rel. 40, Last annotation update)
DE MATRIX METALLOPROTEINASE-20 PRECURSOR (EC 3.4.24.-) (MMP-20) (ENAMEL
DE MMP20).
DE DE MMP20.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SVJ;
RC MEDLINE=20079167; PubMed=10610728;
RA Caterina J., Shi J., Krakora S., Bartlett J.D., Engler J.A.,
RA Kozak C.A., Birkedal-Hansen H.,
RA "Isolation, characterization, and chromosomal location of the mouse
RA enamelysin gene.";
RA Genomics 52:308-311(1999).
CC -1- FUNCTION: DEGRADATES ENAMELIN, THE MAJOR PROTEIN COMPONENT OF THE
CC ENAMEL MATRIX. TWO OF THE MACROMOLECULES CHARACTERISING THE
CC CARLAGEC EXTRACELLULAR MATRIX: AGGREGAN AND THE CARTILAGE
CC OLIGOMERIC MATRIX PROTEIN (COMP). MAY PLAY A CENTRAL ROLE IN TOOTH
CC FORMATION (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: CLEAVES AGGREGAN AT THE 360-ASN-1-PHE-361 SITE
CC (BY SIMILARITY).
CC -1- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: SECRETED. ASSOCIATED WITH THE EXTRACELLULAR
CC MATRIX (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: EXPRESSED SPECIFICALLY IN THE ENAMEL ORGAN.
CC -1- PFM: AUTOACTIVATES AT LEAST AT THE 106-ASN-1-TYR-107 SITE (BY
CC SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 HEMOPEXIN-LIKE DOMAIN.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC
CC METALLOPROTEASE) ALSO KNOWN AS MATRININ SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; AF156956; AAF28472.1; -
DR EMBL; AF156947; AAF28472.1; JOINED.
DR EMBL; AF156948; AAF28472.1; JOINED.
DR EMBL; AF156949; AAF28472.1; JOINED.
DR EMBL; AF156950; AAF28472.1; JOINED.
DR EMBL; AF156951; AAF28472.1; JOINED.
DR EMBL; AF156952; AAF28472.1; JOINED.
DR EMBL; AF156953; AAF28472.1; JOINED.
DR EMBL; AF156954; AAF28472.1; JOINED.
DR EMBL; AF156955; AAF28472.1; JOINED.
DR EMBL; AF155933; AAF28470.1; -
DR MEROPS; M10.019; -
MGD; MGI:1353466; Mmp20.
PROSITE; PS00546; CYSTEINE_SWITCH; 1.
PROSITE; PS00142; ZINC_PROTEASE; 1.
DR PROSITE; PS00024; HEMOPEXIN; FALSE.NEG.
KW Hydrolase; Metalloprotease; Zinc; Calcium; Zymogen; Signal;
KW Extracellular matrix.
FT SIGNAL 1
FT PROPEP 22 106 BY SIMILARITY.
FT CHAIN 107 482 MATRIX METALLOPROTEINASE-20.
FT DOMAIN 292 482 HEMOPEXIN-LIKE.
FT SITE 99 99 CYSTEINE_SWITCH (POTENTIAL).
FT METAL 225 225 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 226 226 BY SIMILARITY.
FT METAL 229 229 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 235 235 ZINC (CATALYTIC) (BY SIMILARITY).
FT DISULFID 295 482 BY SIMILARITY.
SQ SEQUENCE 482 AA; 54373 MW; 366149FAF2BDC8BB CRC64;

Query Match 75.6%; Score 31; DB 1; Length 482;
Best Local Similarity 100.0%; Pred. NO. 57;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 MYPTY 6
Db 243 MYPTY 247

Search completed: September 4, 2001, 15:54:10
Job time: 307 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2001, 15:53:36 ; Search time 80.44 seconds
(without alignments)
13.158 Million cell updates/sec

Title: US-09-630-345-6
Perfect score: 41
Sequence: 1 MYPTYLK 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues
Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTEMBL16.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_unclassified.*
13: sp_invertebrate.*
14: sp_virus.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35	85.4	1088	4 Q9H2T7	Q9H2T7 homo sapien
2	33	80.5	269	1 Q27159	Q27159 methanobact
3	33	80.5	1069	10 Q9LTZ8	Q9LTZ8 arabidopsis
4	32	78.0	252	2 Q9L875	Q9L875 haemophilus
5	32	78.0	553	11 Q70577	Q70577 mus musculu
6	32	78.0	555	11 Q9R0W2	Q9R0W2 rattus norv
7	32	78.0	593	11 P97558	P97558 rattus norv
8	32	78.0	593	11 P70485	P70485 rattus norv
9	32	78.0	986	4 Q94858	Q94858 homo sapien
10	32	78.0	1346	4 Q9UIZ3	Q9UIZ3 homo sapien
11	31	75.6	40	14 Q9IWZ8	Q9IWZ8 human coxa
12	31	75.6	40	14 Q9IWZ7	Q9IWZ7 human coxa
13	31	75.6	50	14 Q66796	Q66796 human echov
14	31	75.6	53	6 Q9XS83	Q9XS83 equus cabal
15	31	75.6	198	14 Q9Q907	Q9Q907 shope fibro
16	31	75.6	261	4 Q9NRE1	Q9NRE1 homo sapien
17	31	75.6	261	4 Q9NRE7	Q9NRE7 homo sapien
18	31	75.6	261	4 Q9GZS2	Q9GZS2 homo sapien
19	31	75.6	262	5 Q19918	Q19918 caenorhabdi

Q9TV55 sus scrofa
Q94467 dictyosteli
Q9VJ65 drosophila
Q51888 prevotella
Q98H5 arabidopsis
Q98A2 arabidopsis
Q9H4D2 homo sapien
Q9VPA2 drosophila
Q9857 cynops pyrr
Q96640 bovine papi
Q18426 caenorhabdi
Q18198 caenorhabdi
Q29350 archaeoglob
Q9TXN1 caenorhabdi
Q9ESN3 mus musculu
Q88445 swine vesic
Q91919 human coxa
Q02173 caenorhabdi
Q91401 pseudomonas
Q98534 paramecium
Q9VUZ1 drosophila
Q9V5Y8 drosophila
Q9QB94 yaba monkey
Q48871 lactobacilli
Q9DHM6 yaba-like d
Q9PIA0 campylobact

ALIGNMENTS

RESULT 1

Q9H2T7 PRELIMINARY; PRT; 1088 AA.
AC Q9H2T7
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE RANBP17.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20568271; PubMed=11024021;
RA Kutay U., Hartmann E., Treichel N., Calado A., Carmo-Fonseca M.,
RA Prehn S., Kraft R., Gorlich D., Bischoff F.R.;
RT Identification of Two Novel RANGTP-binding Proteins Belonging to the
RL J. Biol. Chem. 275:40163-40168(2000).
DR EMBL: AF222747; AAG44255.1; -
SQ SEQUENCE 1088 AA; 124374 MW; A95D5599388EAEFB CRC64;

Query Match 85.4%; Score 35; DB 4; Length 1088;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYL 7
| | | | |
DB 744 MYPTYL 749

RESULT 2

Q27159 PRELIMINARY; PRT; 269 AA.
ID Q27159
AC Q27159;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE HYPOTHETICAL 31.5 KDA PROTEIN.
GN MTH1087.

OS Methanobacterium thermoautotrophicum.
 OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
 OC Methanothermobacter.
 OX NCBI_TaxID=145262;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-DELTA H;
 RX MEDLINE=98037514; PubMed=9371463;
 RA Smith D.R., Doucette-Stamm L.A., DeLoughery C., Lee H.-M., DuBois J.,
 RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
 RA Harrison D., Hoang L., Keagle P., Lum W., Pothier B., Qiu D.,
 RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
 RA Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
 RA McDougall S., Shimer G., Goval A., Pietrowski S., Church G.M.,
 RA Daniels C.J., Mo J.-I., Rice P., Nolling J., Reeve J.N.;
 RT Complete genome sequence of Methanobacterium thermoautotrophicum
 RT deltaH: functional analysis and comparative genomics.";
 RL J. Bacteriol 179:7135-7155(1997).
 DR EMBL; AE000879; AA85576.1; -;
 DR EMBL; AE000879; AA85576.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 269 AA; 31513 MW; 4B7D64EBDC6B8EA7 CRC64;

Query Match 80.5%; Score 33; DB 1; Length 269;
 Best Local Similarity 83.3%; Pred. No. 52;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 YPTVILK 8
 |||||
 DB 175 YPTVILK 180

RESULT 3
 ID Q9LTZ8 PRELIMINARY; PRT: 1069 AA.
 AC Q9LTZ8;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
 DE SIMILARITY TO NEGATIVE REGULATOR OF VESICLE FORMATION.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eucosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-COLUMBIA;
 RX Nakamura Y.;
 RA Nakamura Y.;
 RA MEDLINE=20277480; PubMed=10819329;
 RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
 RT features of the regions of 4,504,864 bp covered by sixty P1 and TAC
 RT clones.";
 RL DNA Res. 7:131-135(2000).
 DR EMBL; AB024028; BAA95706.2; -;
 DR EMBL; AP000381; BAA95706.2; JOINED.
 DR InterPro; IPR000379; -;
 DR InterPro; IPR000734; -;
 DR PROSITE; PS00120; LIPASE_SER; UNKNOWN_1.
 SQ SEQUENCE 1069 AA; 118974 MW; 0BA3545568EC6B44 CRC64;

Query Match 80.5%; Score 33; DB 10; Length 1069;
 Best Local Similarity 83.3%; Pred. No. 2,3e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTVL 7
 |||||

DB 71 MYPTVI 76

RESULT 4
 ID Q9L875 PRELIMINARY; PRT: 252 AA.
 AC Q9L875;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
 DE BETA 1-4 GLUCOSYLTRANSFERASE LGTF.
 GN LGTF.
 OS Haemophilus ducreyi.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Haemophilus
 OX NCBI_TaxID=730;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=35000;
 RX MEDLINE=20278119; PubMed=10816485;
 RA Fillard M.J., Gibson B.W., Schilling B., Sun S., Munson R.S. Jr.,
 RA Campagnari A.A.;
 RT "Construction and Characterization of Haemophilus ducreyi
 RT Lipooligosaccharide (LOS) Mutants Defective in Expression of
 RT Heptosyltransferase III and beta1,4-Glucosyltransferase;
 RT Identification of LOS Glycoforms Containing Lactosamine Repeats.";
 RL Infect. Immun. 68:3352-3361(2000).
 DR EMBL; AF215936; AAF72876.1; -;
 DR InterPro; IPR001173; -;
 DR Pfam; PF00535; Glycos_transf_2; 1.
 KW Transferase.
 SQ SEQUENCE 252 AA; 28736 MW; 20297CB79C070AD4 CRC64;

Query Match 78.0%; Score 32; DB 2; Length 252;
 Best Local Similarity 83.3%; Pred. No. 78;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTVL 7
 :|||||
 DB 134 LYPTVL 139

RESULT 5
 ID O70577 PRELIMINARY; PRT: 553 AA.
 AC O70577;
 DT 01-AUG-1998 (TrEMBLrel. 07, Created)
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
 DE ORGANIC CATION TRANSPORTER 2.
 GN SLC22A2 OR OCT2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129; TISSUE=LIVER;
 RA Mooslehner K.A., Allen N.D.;
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -!- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
 DR EMBL; AJ006036; CAA06827.1; -;
 DR MGD; MGI:1335072; SLC22a2.
 DR InterPro; IPR001066; -;
 DR Pfam; PF00083; sugar_tr; 1.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_2.
 KW Transmembrane.
 SQ SEQUENCE 553 AA; 61830 MW; 1E9744F0D5415483 CRC64;

Query Match 78.0%; Score 32; DB 11; Length 553;

Best Local Similarity 57.1%; Pred. No. 1.8e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 4; Conservative 3;

QY 2 MYPTYLK 8
:||||:
Db 457 LYPTIIR 463

RESULT 6

Q9R0W2 PRELIMINARY; PRT; 555 AA.
AC Q9R0W2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE ORGANIC CATION TRANSPORTER OCT2R.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
[1]
SEQUENCE FROM N.A.
RC TISSUE=KIDNEY;
RX MEDLINE=99316020; PubMed=10385678;
RA Gruendemann D., Lieblich G., Kiefer N., Koester S., Schoemig E.;
RT "Selective substrates for non-neuronal monoamine transporters."; Mol. Pharmacol. 56:1-10(1999).
DR EMBL; Y13154; CAB52215.1; -;
DR InterPro: IPR001066; -;
DR Pfam: PF00083; sugar_tr; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_2.
SQ SEQUENCE 555 AA; 62342 MW; 29521969AEIAC206 CRC64;

Query Match 78.0%; Score 32; DB 11; Length 555;

Best Local Similarity 57.1%; Pred. No. 1.8e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 4; Conservative 3;

QY 2 MYPTYLK 8
:||||:
Db 457 LYPTIIR 463

RESULT 7

P97558 PRELIMINARY; PRT; 593 AA.
AC P97558;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE ORGANIC CATION TRANSPORTER.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
[1]
SEQUENCE FROM N.A.
RC TISSUE=KIDNEY;
RA Gorboulev V.G., Koepsell H.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
DR EMBL; X98334; CAA66979.1; -;
DR InterPro: IPR001066; -;
DR Pfam; PF00083; sugar_tr; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_2.
KW Ion transport; Transmembrane.
SQ SEQUENCE 593 AA; 66100 MW; 36C3E1B5DC057790 CRC64;

Query Match 78.0%; Score 32; DB 11; Length 593;

Best Local Similarity 57.1%; Pred. No. 2e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 4; Conservative 3;

QY 2 MYPTYLK 8
:||||:
Db 457 LYPTIIR 463

RESULT 8

P70485 PRELIMINARY; PRT; 593 AA.
AC P70485;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE ORGANIC CATION TRANSPORTER OCT2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
[1]
SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=KIDNEY;
RX MEDLINE=96295517; PubMed=8702418;
RA Okuda M., saito H., Urakami Y., Takano M., Inui K.;
RT "cDNA cloning and functional expression of a novel rat kidney organic cation transporter, OCT2."; Biochem. Biophys. Res. Commun. 224:500-507(1996).
RL Biochem. Biophys. Res. Commun. 224:500-507(1996).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
DR EMBL; D83044; BAA11754.1; -;
DR InterPro: IPR001066; -;
DR Pfam; PF00083; sugar_tr; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_2.
KW Transmembrane.
SQ SEQUENCE 593 AA; 66080 MW; 36C1044E0C04B443 CRC64;

Query Match 78.0%; Score 32; DB 11; Length 593;

Best Local Similarity 57.1%; Pred. No. 2e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 4; Conservative 3;

QY 2 MYPTYLK 8
:||||:
Db 457 LYPTIIR 463

RESULT 9

O94858 PRELIMINARY; PRT; 986 AA.
ID O94858;
AC O94858;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE KIAA0758 PROTEIN (FRAGMENT).
GN KIAA0758.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
RP SEQUENCE=BRIN;
RC MEDLINE=99087487; PubMed=9872452;
RX Nagase T., Ishikawa K., Suyama M., Kikuno R., Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RA "Prediction of the coding sequences of unidentified human genes. XI. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro."; DNA Res. 5:277-286(1998).
RL DNA Res. 5:277-286(1998).
DR EMBL; AB018301; BAA34478.1; -;
DR InterPro: IPR000203; -;
DR PROSITE; PS00083; -;
DR InterPro: IPR000832; -;
DR Pfam; PF00002; 7tm2; 1.
DR Pfam; PF01825; GPS; 1.

DR PRINTS; PR00249; GPCRSECRETIN.
DR SMART; SM00303; GPS; 1.
FT NON_TER 1
SQ SEQUENCE 986 AA; 108720 MW; EF4875B03BAA0E7D CRC64;

Query Match 78.0%; Score 32; DB 4; Length 986;
Best Local Similarity 85.7%; Pred. No. 3.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MYPTLYK 8
| | | | |
Db 384 MLPTLYK 390

RESULT 10
Q9UI23 PRELIMINARY; PRT; 1346 AA.
AC Q9UI23;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE DJ365012.1 (KIAA0758 PROTEIN).
GN DJ365012.1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Sehra H.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX DOMAIN.
DR EMBL; AL096772; CAB61578.1; -.
DR InterPro; IPR000082; -.
DR InterPro; IPR000203; -.
DR InterPro; IPR000561; -.
DR InterPro; IPR000832; -.
DR InterPro; IPR003006; -.
DR Pfam; PF00002; 7tm_2; 1.
DR Pfam; PF00047; 1q_2; 1.
DR Pfam; PF01825; GPS; 1.
DR PRINTS; PR00249; GPCRSECRETIN.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR SMART; SM00300; SEA; 1; 149456 MW; 72A9D02B08218A60 CRC64;
SQ SEQUENCE 1346 AA; 149456 MW; 72A9D02B08218A60 CRC64;

Query Match 78.0%; Score 32; DB 4; Length 1346;
Best Local Similarity 85.7%; Pred. No. 4.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 MYPTLYK 8
| | | | |
Db 744 MLPTLYK 750

RESULT 11
Q9IWZ8 PRELIMINARY; PRT; 40 AA.
AC Q9IWZ8;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE VP2 PROTEIN (FRAGMENT).
GN VP2.
OS Human coxsackievirus B5.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCBI_TaxID=12074;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=R8597;
RA Casas I., Palacios G.F., Trallero G., Cisterna D., Freire M.C.,
RT Tenorio A.;
RT "Molecular characterization of human enteroviruses in clinical samples
by three different RT nested PCR assays and direct sequencing of
RT amplified products";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF252193; AAF71330.1; -.
DR InterPro; IPR001676; -.
DR Pfam; PF00073; rhv; 1.
DR NON_TER 1
FT NON_TER 40
SQ SEQUENCE 40 AA; 4388 MW; 7F998A9865EF9F67 CRC64;

Query Match 75.6%; Score 31; DB 14; Length 40;
Best Local Similarity 71.4%; Pred. No. 17;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTLYK 8
| | | | |
Db 26 VWPTLYK 32

RESULT 12
Q9IWZ7 PRELIMINARY; PRT; 40 AA.
AC Q9IWZ7;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE VP2 PROTEIN (FRAGMENT).
GN VP2.
OS Human coxsackievirus B5.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCBI_TaxID=12074;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R8497;
RA Casas I., Palacios G.F., Trallero G., Cisterna D., Freire M.C.,
RT Tenorio A.;
RT "Molecular characterization of human enteroviruses in clinical samples
by three different RT nested PCR assays and direct sequencing of
RT amplified products";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF252194; AAF71331.1; -.
DR InterPro; IPR001676; -.
DR Pfam; PF00073; rhv; 1.
DR NON_TER 1
FT NON_TER 40
SQ SEQUENCE 40 AA; 4388 MW; 7F998A9865EF9F67 CRC64;

Query Match 75.6%; Score 31; DB 14; Length 40;
Best Local Similarity 71.4%; Pred. No. 17;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTLYK 8
| | | | |
Db 26 VWPTLYK 32

RESULT 13
Q66796 PRELIMINARY; PRT; 50 AA.
AC Q66796;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE CAPSID PROTEIN VP2 (FRAGMENT).
GN VP2.
OS Human echovirus 20.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCBI_TaxID=47508;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JV-1;
RX MEDLINE=96203959; PubMed=8627260;
RA Huttunen P., Santti J., Pulli T., Hyypia T.;
RT "The major echovirus group is genetically coherent and related to
RT coxsackie B viruses.";
RL J. Gen. Virol. 77:715-725(1996).
DR EMBL; X89549; CAA61727.1; -;
DR InterPro; IPR001676; -;
DR Pfam; PF00073; rnv; 1.
FT NON_TER 1
FT NON_TER 50
SQ SEQUENCE 50 AA; 5448 MW; BBFC58B2511FC2E7 CRC64;

Query Match 75.6%; Score 31; DB 14; Length 50;
Best Local Similarity 71.4%; Pred. No. 22;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYLK 8
DB 31 VWPYTLK 37

RESULT 14
O9XS83 PRELIMINARY; PRT; 53 AA.
ID Q9XS83
AC Q9XS83;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE TUBBY PROTEIN (FRAGMENT).
GN TUB.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99160468; PubMed=10051323;
RA Caetano A.R., Pomp D., Murray J.D., Bowling A.T.;
RT "Comparative mapping of 18 equine type I genes assigned by somatic
RT cell hybrid analysis.";
Mamm. Genome 10:271-276(1999).
EMBL; AF097580; RAD25983.1; -;
HSSP; P50586; 1C8Z.
DR InterPro; IPR000007; -;
DR Pfam; PF01167; Tub; 1.
FT NON_TER 1
FT NON_TER 53
SQ SEQUENCE 53 AA; 6244 MW; 13A9ED8C69500D3E CRC64;

Query Match 75.6%; Score 31; DB 6; Length 53;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPT 6
DB 16 MYPT 20

RESULT 15
Q9Q907 PRELIMINARY; PRT; 198 AA.
ID Q9Q907
AC Q9Q907;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)

DE GP064R.
GN S064R.
OS Shope fibroma virus (strain Kasza) (SFV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Leporipoxvirus.
OX NCBI_TaxID=10272;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KASZA;
RX MEDLINE=84165064; PubMed=6323741;
RA Delange A.M., Macaulay C., Block W., Mueller T., McFadden G.;
RT "Tumorigenic poxviruses: construction of the composite physical map of
RT the Shope fibroma virus genome.";
RL J. Virol. 50:408-416(1984).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=KASZA;
RX MEDLINE=92074222; PubMed=1660196;
RA Strayer D.S., Jerng H.H., O'Connor K.;
RT "Sequence and analysis of a portion of the genomes of Shope fibroma
RT virus and malignant rabbit fibroma virus that is important for viral
RT replication in lymphocytes.";
RL Virology 185:585-595(1991).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=KASZA;
RX MEDLINE=20032074; PubMed=10562495;
RA Willer D.O., McFadden G., Evans D.H.;
RT "The complete genome sequence of shope (Rabbit) fibroma virus.";
RL Virology 264:319-343(1999).
DR EMBL; AF170722; AAF17946.1; -;
SQ SEQUENCE 198 AA; 23282 MW; 9D181052C1ED42C5 CRC64;

Query Match 75.6%; Score 31; DB 14; Length 198;
Best Local Similarity 83.3%; Pred. No. 96;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MYPTYL 7
DB 110 MYPTFL 115

Search completed: September 4, 2001, 15:53:37
Job time: 290 sec

